WHEN INPUTS ARE OUTPUTS:
THE CASE OF GRADUATE STUDENT INSTRUCTORS[*]

Eric P. Bettinger
Stanford University
and NBER

Bridget Terry Long
Harvard Graduate
School of Education
and NBER

Eric S. Taylor
Harvard Graduate
School of Education

October 2015

We examine graduate student teaching as an input to two production processes: the education of undergraduates and the development of graduate students themselves. Using fluctuations in full-time faculty availability as an instrument, we find undergraduates are more likely to major in a subject if their first course in the subject was taught by a graduate student, a result opposite of estimates that ignore selection. Additionally, graduate students who teach more frequently graduate earlier and are more likely to subsequently be employed by a college or university.

JEL Classifications: I2, J2

I. Introduction

There are few cases in economics where inputs into the production function are also outputs in the production function. However, such is the case in the production of doctoral students. Doctoral students are essential inputs to significant outputs in higher education. They provide research support, generate peer effects on each other, and often teach undergraduate students. Doctoral students are also an important output in education. Graduate programs define and often rank themselves by the quality and quantity of doctoral students they produce. Moreover, among those who go on to work in higher education, the productivity of graduate students in their subsequent careers is another important "output" of doctoral programs in that it increases the prominence of the doctoral-granting organization, increases the status of the doctoral adviser, and contributes to the subsequent education of new scholars in the field.

To date, there is very little research on the productivity of graduate students and on the relationships between their graduate activities and subsequent careers, particularly with regards to their teaching responsibilities while enrolled in school. What research there is with regards to the teaching function of graduate students, however, suggests that undergraduates that have a graduate student as their instructor, particularly non-native English speaking graduate students, suffer worse outcomes than comparable students that have faculty instructors (e.g., Borjas 2000). Moreover, there exists no evidence on how graduate students' teaching experiences contribute to their subsequent careers. [1]

---

[1] In an earlier NBER working paper (Bettinger and Long 2004), we reported on the productivity of graduate students and adjuncts in undergraduate teaching. That analysis was the precursor to two separate analyses. Given the different nature of adjunct usage (i.e. largely focused on cost considerations) from graduate student usage, we focused the final version solely on adjuncts (Bettinger and Long 2010). In this paper, we return to the issue of graduate students with a different sample, a revised empirical strategy, and additional information on tracking academic and job market outcomes of the graduate student instructors themselves. Given the distinction in the labor market considerations between graduate student employment and adjuncts, the mechanisms and

This paper examines graduate student teaching as an input to two production processes: the education of undergraduates and the development of graduate students themselves. As such we attempt to answer two related research questions. First, we quantify the effect of graduate student instructors on the academic outcomes of the undergraduate students they teach. As outcomes we examine students' choice of major, course taking, and credits earned. While graduate students participate in several facets of knowledge production and instruction, this is perhaps the most common avenue for their participation. In this paper, we are distinguishing between the roles of being a teaching assistant (i.e., course support) versus serving as the primary teacher of a course (i.e., the instructor); our analysis focuses on the latter. About 46 percent of undergraduate students at four-year colleges take courses that have graduate students as their primary instructors,[2] and in 2000, over 70 percent of graduate students had some teaching responsibility. At issue is whether using graduate students as instructors relative to other possible types of instructors (faculty or adjuncts) is a good way to allocate resources as measured by the outcomes of undergraduate students.

Second, we examine how teaching experience contributes to graduate students' subsequent academic careers. One of the justifications for using graduate students as instructors is that it provides essential training for their subsequent careers. We measure the relationship between teaching experiences and the subsequent academic outcomes and career choices of the graduate student. We examine the likelihood the graduate student completes their doctorate and is later employed at a college or university.

For our analysis, we use administrative data from Ohio's public universities. To answer the first question, we track the outcomes of undergraduate students who

institutional considerations behind our results may shed light on the underlying economic and policy considerations in graduate student education.

[2] Authors' calculations based on BPS 2003/2009 Sample.

initially enrolled in college in the fall of 1998 or 1999. We identify the first course (e.g. Introductory Economics) that an undergraduate student took in a specific department (e.g. Economics) and observe whether these courses were taught by graduate students as opposed to full-time or part-time faculty. Using exogenous variation in faculty availability, we estimate the causal impact of graduate student instructors on undergraduate students' subsequent course-taking behavior.

To answer the second question, we study students who began their graduate program in 1998 or 1999, and who subsequently taught at least one course. We link university administrative data with data from the Ohio Department of Jobs and Family Services (ODJFS). ODJFS tracks employment outcomes throughout the state, and these data allow us to measure graduate students' subsequent earnings and location of employment, so long as they stay within the state. We use the combined data to relate the number of teaching opportunities graduate students experience to their subsequent academic and professional outcomes (within six years of starting graduate work). Although our data are limited to observing subsequent behavior in Ohio alone, six years after the start of their graduate program, we find 82.5 percent of graduate students in either our employment data (52.6 percent) or university administrative data (29.9 percent, Table 4). Therefore, we believe we are observing subsequent outcomes for the vast majority of students. Moreover, while limitations remain, this is the first large-scale study to shed light on the relationship between teaching and outcomes after graduate school.

Our research questions are fraught with concerns about selection. Graduate students are unlikely to randomly choose to teach a course, and other work suggests that undergraduates may actively shy away from taking courses taught by certain kinds of instructors (Bettinger and Long, 2010). We use multiple strategies to deal with these issues. To address the fact that students may sort non-randomly across courses, we present the results using (a) course-by-department fixed effects, which limits the identifying variation to students who took the same course at the same

institution but from different instructor types, and (b) course-by-department-by-term fixed effects, which limits identification to students who took the same course at the same time but from different instructors. Additionally, we estimate the results using an instrumental variables approach that capitalizes on the natural fluctuations from term to term in the faculty available to teach. In terms of whether a graduate student chooses to teach, we use department fixed effects to look within departments. However, there may be residual selection bias and so we caution about the interpretation of the results. Still, using a variety of assumptions about the size and direction of this bias, we still believe these results shed light on this understudied issue.

Our results suggest that graduate students are effective instructors relative to faculty members—at least as judged by the measures of their student's subsequent academic progress we can observe. Undergraduates taught by graduate students in a given subject are more likely to subsequently major in the subject compared to their peers who take the same course from full-time faculty. However, we find no statistically significant differences in the number of subsequent credits earned in the subject. Given that we use random variation in students' exposure to graduate students, we argue that our estimates suggest causal relationships. Second, graduate students who teach more frequently are more likely to complete their doctoral degree in a timely manner and more likely to be employed subsequently by a college or university. Regardless of whether we interpret these as causal or selection, the results suggest graduate student teaching benefits the sector. If our results only reflect selection, that selection, at least in our sample, identifies and incorporates effective future faculty into undergraduate production. By contrast, if our results reflect causality, then they suggest that undergraduate instructional experiences positively impact the short-run, academic job prospects of graduate students.

In the next section, we discuss existing academic literature on graduate student teaching. In section II we describe our empirical strategies and the Ohio sample and setting. Section III presents and discusses the results, and Section IV concludes.

II. Existing Literature

The input-output duality of graduate student teaching has long been recognized in American colleges and universities. During the rapid growth of graduate student teaching after World War II, one presenter at the 1966 annual meeting of the American Council on Education stated that "universities, in their need to staff elementary courses, have thoroughly abused…the chief means of training prospective college teachers—the teaching assistantship" (Wise 1967). Scholars in education, economics, and sociology have contributed to an academic literature on the topic which has largely followed the two research questions at the heart of this paper.

The first set of results has focused on the effectiveness of graduate students as instructors. Ex-ante, it is unclear what we should expect about the quality of graduate student teaching. On the one hand, teaching experience can be a meritocracy of sorts, and department administrators may attempt to place the "best" teaching graduate students in prime teaching slots. Yet some researchers have highlighted the poor working conditions of graduate students (e.g. McLeod and Schwarzbach 1993). These researchers document how low wages, large workloads, and working without training or guidance may actually generate negative effects (Koehnecke 1991). In fact, these stresses coupled with the underlying trend toward increased reliance on graduate student instructors have led to recent efforts to strengthen graduate student teacher unions at universities across the nation (Mattson 2000, Vaughn 1998, Sharnoff 1993).

Most of the empirical literature of graduate student instruction quality has focused on the effects of foreign graduate students on undergraduate students' outcomes. Borjas (2000) studied graduate student instruction in economics at a large public institution in the United States. In that setting, undergraduates taught by foreign-born teaching assistants received lower final grades by about 0.2 grade points relative to US-born graduate student instructors. However, the negative effect was muted for foreign-born undergraduates and when the graduate student was given higher class-preparedness ratings by the undergraduates. Borjas' (2000) results are consistent with preceding findings in economics instruction (Watts and Lynch 1989), though in other subjects the limited evidence is mixed (Jacobs and Friedman 1988, Norris 1991). Subsequently, Fleisher, Hashimoto, and Weinberg (2002) found no differential, and possibly positive, effects of foreign graduate student instructors when they received training and completed at least one year of their own graduate study prior to taking-on teaching responsibilities.

Beyond comparisons of foreign and native graduate students, empirical evidence on instructional effectiveness in the college setting is scarce. This is in stark contrast to the American K-12 setting, for which there are numerous studies estimating the range of outcomes students with different teachers have and how this variance is related to teacher characteristics such as training level and certification. There are two studies that document the variation in instructor effectiveness. For example, using variation in instructor quality at the US Air Force Academy and benefitting from the random assignment of cadets to course sections, Carrell and West (2010) estimate the standard deviation in instructor effectiveness as approximately 0.052 student standard deviations on a common end-of-course exam. Hoffman and Oreopoulos (2009) find similar estimates for common exam courses (0.055 standard deviations), but the standard deviation is larger for all courses. These estimates are somewhat smaller than the evidence on variability in K-12 teacher effectiveness (Hanushek and Rivkin 2010).

Beyond the general variance of outcomes across instructors, the above two studies also look at the effects of particular instructor characteristics. Carrell and West (2010) find that students assigned less-experienced instructors and instructors without a terminal degree (two correlated characteristics) score higher in the short-run but perform less-well in subsequent classes that build on content from the first. Hoffman and Oreopoulos (2009) find that part-time instructors, which include both adjuncts and graduate students in their sample, have no differential effect on undergraduate outcomes compared to full-time faculty. Another study focuses on the relative effects of full-time versus part-time faculty. The authors find small positive effects for adjuncts compared to full-time faculty in terms of students' subsequent interest in the subjects; however, the effects vary by subject (Bettinger and Long, 2010). However, other work using the same data find that the percentage of instructors a freshman student has that are adjuncts during their first year is negatively related to college persistence (Bettinger and Long, 2006).

Our primary outcome of interest is students' subsequent engagement in a course as measured by decisions to major, decisions to take additional courses, and the total number of subsequent classes that students took. Our data limit us from being able to track performance in the current or any subsequent course. However prior literature on teaching quality in K-12 frequently equates good teaching with subsequent engagement in courses (e.g. Skinner and Belmont 1993, Reeve and Jang 2006), and similar connections exist in higher education (e.g. Coates 2005). While subsequent engagement may be the result of other graduate student characteristics (i.e. generating interest through entertainment, ease of grading or other mechanisms), the prior literature certainly finds a correlation with overall teaching quality.

A second set of papers focuses on the role of graduate student teaching as an input to their own course of study and development. Researchers have long-argued that graduate student teaching experience is essential to training future

professors (e.g. Smith 2001, Meyers and Prieto 2000, Knotts and Main 1999, Prieto and Altmaier 1994, Slevin 1992). Most existing empirical evidence measures proximate outcomes for doctoral students, including degree completion and time to completion, which are a present concern for American higher education (Bowen and Rudenstine 1992). A few studies have found that graduate students who receive teaching assistantships, as compared to other forms of financial aid, are less likely to complete their degree and those who do finish take longer to do so (Ehrenberg and Mavros 1995, Seagram, Gould, and Pyke 1998). These effects may differ from discipline to discipline, and students with teaching assistantships do better than those who have no institutional support (Bair and Haworth 2004).

By contrast, there is little empirical evidence on how teaching during one's time as a graduate student affects long-run outcomes, including subsequent labor market prospects, research productivity, and other academic and professional outcomes. In this paper, we begin to address this hole in the empirical literature using data which track graduate students' employment after degree completion.

III. Empirical Strategy

Outcomes for Undergraduate Students

Our first objective is to identify the effects of taking a class taught by a graduate student on undergraduates' future course-taking behavior and choice of major. We focus on variation in the instructors of undergraduate students' initial, often introductory, course in a given subject because that early experience may affect long-term interest and success in the subject area.

Equation 1 describes the core of our approach: predict different outcomes, $y_{ijkl(t+1)}$, for undergraduate student i in subject k at campus j at time $t+1$ as a function of whether they were taught by a graduate student in their introduction to the subject at time t.

$$(1) \quad y_{ijk(t+1)} = \alpha + \beta GradStu_{ijkt} + \varphi Adjunct_{ijkt} + X_i\gamma + Z_{ik}\delta + \pi_t +$$
$$\tau_k + \sigma_j + \varepsilon_{ijk(t+1)}$$

Operationally, $GradStu_{ijkt}$, the explanatory variable of interest, will be equal to one if the student took her first course in the subject from a graduate student and zero if she took the course from someone other than a graduate student. If she took multiple courses in a given subject during the first term of exposure $GradStu_{ijkt}$ will equal the proportion of courses taught by a graduate student.

We use this structure to study three different outcomes: (i) whether student i took any additional courses in subject k beyond time t, (ii) the total number of credit hours taken in subject k beyond time t, and (iii) whether student i majored in subject k. Since undergraduate students' course and major choices often lay the groundwork for future careers, graduate student instruction may affect long-run labor market outcomes for both individuals and at a more macro-level, through these nearer-term decisions. These mechanisms could, for instance, change the composition of available types of labor if undergraduates sort among disciplines as a result of exposure to high- or low-quality graduate student teaching.

In addition to the explanatory variable of interest, we control for a vector of student characteristics, X_i, which includes demographics (gender, race/ethnicity), initial ability (measured by ACT score), state of residence, and year of college entry (an indicator for being in the fall 1999 versus 1998 cohort). We also account for each students' expressed interest in and exposure to subject k with the vector Z_{ik}, which captures whether the subject is the student's intended major and the number of credits hours taken in the subject during the first term of exposure. Our specification also includes fixed effects for the term during which first exposure occurred, π_t, the subject, τ_k, and campus, σ_j. Finally, we also include a control for exposure to adjunct instructors in initial courses, $Adjunct_{ijkt}$, which we define in

the same way as $GradStu_{ijkt}$. Separating out $Adjunct_{ijkt}$ clarifies the interpretation of our coefficient of interest so that β only measures the effect of graduate student instructors relative to full-time faculty.[3]

Because students are observed in multiple subjects, we address within-student correlation by clustering the standard errors at the student level. Additionally, since the treatment of interest (which type of instructor the student is exposed to) varies by section, we also report standard errors clustered at the section-by-course level.

To interpret our estimate of β causally requires the assumption that undergraduates taught by graduate students are not systematically different from other undergraduates, at least conditional on the included variables in equation 1. A likely violation is the non-random way in which students and instructors are sorted into course sections. Such sorting may occur across courses within departments, across terms, or across sections of a course taught in the same term. For example, students may have prior beliefs about the relative effectiveness of graduate student instructors and so may try purposely to avoid them and enroll in course taught by full-time faculty members. Thus we report estimates which develop our core approach (equation 1) in three ways.

First, we present results which add course-by-department fixed effects; we define departments as the interaction of subject and campus. This limits the identifying variation to students who took the same course at the same institution, but from different instructor types. The restriction addresses differences in the distribution of instructor types across departments and differences in how departments assign courses to different instructor types. Still, student decisions about which term to take a course may be influenced by expectations of future

---

[3] The effects of adjunct instructors relative to full-time faculty are studied in a similar framework with the same data by Bettinger and Long (2010).

instruction type options. Thus, second, we report estimates with course-by-department-by-term fixed effects, limiting identification to students who took the same course at the same time but from different instructors. To the extent students are randomly assigned to course sections in a given term, these course-by-department-by-term fixed effects estimates will be unbiased. But even this assumption is potentially too strong. For example, full-time faculty may be more likely to teach honors sections or daytime sections with regular or late-hours sections left to graduate students. Students planning to major in a given subject may actively seek-out full-time faculty per se or honors sections.

To address the remaining selection we employ a third strategy: an instrumental variables approach to limit identification to variation in instructor type that is uncorrelated with prior, unobserved student interest in the subject. Specifically, we instrument for instructor type, identified by the variables $GradStu_{ijkt}$ and $Adjunct_{ijkt}$ in Equation 1, using each department's term-by-term variation in the proportion of courses taught by professors of different rank. Departments face natural fluctuations from term to term in the faculty available to teach as professors are hired, retire, take sabbaticals, and replace courses with research grant funds. Departments also experience temporary fluctuations in the demand for particular courses that translates to more course sections. Since graduate students are one solution to short-run staffing fluctuations, we hypothesize that variation in the proportion of full-time professors from term to term will predict undergraduates' exposure to graduate student instruction. Likewise, this instrument should predict exposure to part-time instructors, and we therefore instrument for adjuncts as well.[4]

The first-stage, described by Equation 2, models the probability of exposure to graduate student instruction as a function of the department-specific variation

---

[4] This is the same approach adopted in Bettinger and Long (2010).

from term to term in the proportion of courses taught by full-time professors relative to a department steady state.

$$
\begin{aligned}
(2) \quad GradStu_{ijkt} = {} & \alpha + \eta AsstProf_{jkt} + \theta AssocProf_{jkt} + \lambda FullProf_{jkt} + \\
& X_i\gamma + Z_{ik}\delta + \pi_t + \tau_k + \sigma_j + \varepsilon_{ijkt}
\end{aligned}
$$

First, for each department k at campus j, we calculate the number of assistant, associate and full professors teaching during time $t,$ where t is the term of a specific school year.[5] To address differences in scale across departments, we normalize these counts by the typical number of full-time professors teaching in the department during the same term.[6] These normalized counts are captured in the variables $AsstProf_{jkt}$, $AssocProf_{jkt}$, and $FullProf_{jkt}$ in Equation (2). Second, we include campus and department fixed effects, σ_j and τ_k respectively, so that we are identifying based on *deviations* from a given department's average reliance on different types of faculty from term to term.

As reported below, deviations in the proportion of faculty teaching in a given term are highly correlated with the probability that an undergraduate took her initial course from a graduate student. However, for our instrumental variables estimates to be unbiased, this distribution of teaching across ranks should not affect undergraduate outcomes directly. Consider, for example, a department in long-run decline where the number of full-time faculty shrinks in successive time periods and replacing them is unlikely to take place. To address short-run course

---

[5] This calculation is made using the transcripts of entering fall 1998 and 1999 freshman cohorts. Given this data limitation, our measures will not count professors whose courses only enrolled students who matriculated before fall 1998.

[6] Specifically, we calculate the number of full-time professors (i.e., the sum of assistant, associate, and full professors) each term each year (i.e., 1998 fall through 2003summer), and then take the average across years for each term: fall, winter, spring, and summer. This term-specific five-year average is the denominator in our normalization.

commitments, the department would likely turn to adjunct and graduate student instructors. In this scenario, students might be reluctant to continue their studies in a collapsing department, even if instructor type has no independent relationship with undergraduate students' future academic trajectories. We are not particularly concerned about this threat, however, as prior analysis (Bettinger and Long, 2010) suggests that such declining departments, representing just three percent of all departments across all schools, are rare in our analysis sample.

Outcomes for Graduate Student Instructors

Our second objective is to explore the relationship between individual graduate students' teaching experiences and their own academic and professional outcomes. One way to interpret this analysis is causal: that the teaching experiences could affect the subsequent outcomes of the graduate students. However, given that graduate students do not randomly choose whether to teach, we acknowledge the likely selection involved. Therefore, we pursue this analysis in part to gain insight into the selection of graduate student *instructors* from among the large pool of graduate students in a department. The limited evidence on graduate student teaching, notably Borjas (2000), suggests that understanding which graduates are teaching is critical to understanding the effects of their teaching on undergraduate students.

We examine this relationship by estimating the probability that graduate students complete a doctoral degree, and separately, gain employment at another college or university in Ohio within six years of beginning their graduate program as a function of their teaching experience during graduate school. Equation 3 describes our approach.

$$(3) \qquad y_{ijk} = \beta TermsTaught_{ijk} + \tau_{kj} + \varepsilon_{ijk}$$

The key predictor variable, $TermsTaught_{ijk}$, is the number of terms graduate student i taught an undergraduate class (within six years of starting graduate school). Our specification also includes fixed effects for department-by-campus τ_{kj}. Standard errors are clustered at the department-by-campus level.

We estimate equation 3 using our sample of graduate students who began their graduate programs in 1998 or 1999, and subsequently taught at least one undergraduate class.[7] Thus $TermsTaught_{ijk}$ cannot be zero, and the correct interpretation of β is how each additional term teaching, conditional on teaching in one term, predicts the outcome of interest. The data available for this study do not include graduate students with zero teaching experience, thus we cannot estimate the margin of teaching versus no teaching experience.

Selection is again a primary concern for any casual interpretation of estimated effects on graduate student outcomes. Notably, when selecting an instructor from among a pool of graduate students, departments may be more likely to select graduate students who are planning to seek a faculty job after graduation; alternatively, they may try to protect those students from teaching so that they can focus on building a research portfolio. From the perspective of the graduate student, the ability to secure teaching opportunities may be a function of interest, motivation, and other latent characteristics that are correlated with academic and professional success. While fixed effects reduce some sources of bias by identifying on differences between students in the same department, at this time we do not have additional empirical strategies to address residual selection bias. As a result, we caution against interpreting these estimates causally; rather, we believe these results help clarify the implications of our estimates for undergraduate students.

---

[7] Specifically, they taught at least one class that included undergraduate students who were first-time freshmen in fall 1998 or 1999.

Data and Study Sample

To estimate the specifications described above, we use administrative data provided by the Ohio Board of Regents (OBR). The undergraduates in our sample are first-time, full-time, traditional-age (18 to 20 years) freshman who enrolled at one of twelve public, four-year colleges and universities in Ohio during the fall of 1998 or the fall of 1999. The graduate students in our sample were enrolled at one of the same twelve institutions and taught undergraduates from one or both of these two cohorts at least once during the undergraduates' college careers.

For undergraduates the data include demographic details, test scores, major, college courses taken, and the instructor(s) for each course. The OBR data track students across campuses within the Ohio public higher education system, both two- and four-year institutions, which enables us to observe students even when they move from one school to another, or withdraw from any one course. Thus, for example, we can link a student's first biology class at Ohio State with a biology major later declared at the University of Cincinnati. Similarly we can track students who took their first biology course from a graduate student, even if the student withdrew from the course and did not receive a grade. Unfortunately, we do not observe outcomes at private institutions and institutions outside Ohio, and as a result, we cannot distinguish students who transfer to a school outside the Ohio public system from students who drop out of college entirely. Previous work by Bettinger (2004) suggests this is not likely to impact our substantive findings since the rate of transfer to out-of-state or private institutions is quite small among our student sample.

Additionally, for most undergraduates the data also include information taken from forms students complete when sitting for the ACT exam. Particularly important for our identification strategy, these ACT data include the student's intended major stated before they matriculate. Given the value of knowing students' major intentions, we limit our undergraduate analysis sample to individuals who

took the ACT. In our data, 89 percent of four-year college students in Ohio take the ACT.[8]

Table 1 reports summary statistics for these undergraduates and their course taking behavior. Consistent with national trends, the sample is more female and less black than Ohio's college-aged population, as shown in the top panel. They are also spread across Ohio's selective and non-selective public institutions which span the state's urban, suburban, and rural areas. Additionally, the proportion of Ohio high school graduates who immediately enroll in college is near the national average (Mortenson 2002).

The bottom panel of Table 1 describes students' experiences during their first course in a given subject and subsequent behavior in that subject.[9] Across all courses, undergraduates are taught by graduate students in about one-fifth of their courses in a subject, and undergraduates take, on average, five additional credits in the same subject within a year of taking an introductory course.

The right hand column of Table 1 focuses on our undergraduate analytic sample. In our analysis of undergraduate outcomes we focus on undergraduates' first course in a given subject, and further limit the sample to only courses where at least 10 percent, but no more than 90 percent, of sections are taught by graduate student instructors. We also exclude remedial or developmental courses because of their unique design and objectives. This restricted sample of undergraduates is not very different in composition from the full student sample shown on the left of Table 1; however, the restricted sample is drawn slightly more from the selective institutions which are more likely to have graduate programs.

---

[8] Observable student characteristics, the probability of exposure to adjunct and graduate instructors, and the outcomes of interest do not differ across samples with and without this restriction.

[9] Our identification of a student's "first course" does not require that the student completed the course. Thus, even if the student withdrew from the course during the term, and did not receive a grade, the course would still be their first course.

Table 2 compares, by subject, the proportion of undergraduates who took their first course from a graduate student instructor calculated for the full sample and for the sample of courses with variation in instructor type. The restriction affects inclusion of social science and science courses more than some others, such as English and math/statistics, and particularly affects professional departments.

Table 3 characterizes the instructors to which the undergraduates in our student sample were exposed. The first two columns of Table 3 describe the faculty who taught any undergraduate: column 1 shows the raw means at the course-level, and column 2 weights by the number of students taught. Approximately 44 percent of students were taught by faculty with a Ph.D., 21 percent by adjuncts, and 14 percent by graduate students. The right side is limited to faculty of courses where the course was the undergraduate's first introduction to a department. Graduate student instructors were much more likely to be women or minority than full-time instructors.

In addition to the basic instructor characteristics summarized in Table 3, the Board of Regents provided additional data on graduate student instructors. The graduate student data include course taking information (similar to the data for undergraduates), longitudinal data on credits earned and GPA, and detailed degree information. Additional data record the details of employment (i.e., employer, earnings) between the last quarter of 1998 and the second quarter of 2006, but only for employment in Ohio.

Table 4 provides mean characteristics for the sample of graduate student instructors in our analysis of graduate student outcomes; students who began their graduate studies at an Ohio public institution in 1998 and 1999. All students in the sample taught undergraduates at least once (within six years of beginning graduate school); 22 percent taught just once, while 26 percent taught in five or more terms. Within six years of starting graduate school, 16 percent had completed a PhD, and 53 percent were employed in Ohio with 15 percent working at an Ohio college or

university (excluding their own graduate institution). Another 30 percent were still enrolled at their university during the sixth year. As column 2 shows, these characteristics are fairly consistent when focusing exclusively on students from academic departments,

IV. Results

OLSs Estimates of the Effects of Graduate Student Instructors

We begin with OLS estimates of how early exposure to graduate student teaching affects future undergraduate interest in the subject. In general, the OLS estimates suggest negative effects although, as previously discussed, these may be influenced by non-random sorting of instructors and students even after the inclusion of different fixed effects. Table 5 reports OLS estimates predicting the total number of subsequent credit hours taken in a given subject with varying fixed-effects approaches to address some forms of selection. Columns 1-4 of Table 5 report outcomes following exposure to graduate student teaching in an initial course in a subject during any term. The naïve estimates in Column 1, which exclude any fixed effects would suggest having a graduate student instructor reduces the number of future credits taken in the same subject. When we account for systematic differences from term-to-term and fixed differences across campuses and subject areas, Column 2, the effect is reduced by about 60 percent. Similarly, when we further control for differences across courses in the same department, Column 3, the effect is again reduced by about half; it approaches zero but is still statistically significantly negative when we cluster at the student level.[10] Finally, if selection patterns differ over time, a course-by-time fixed effect will reduce bias. When we add these last fixed effects in Column 4, the magnitude of the negative effect increases.

---

[10] Standard errors clustered at the student level are reported in parentheses. Standard errors clustered at the course section level are reported in brackets below the student level errors.

The differences in estimates across Columns 1 through 4 of Table 5 reinforce the importance of considering how departments, courses, and sections of the same course differ in terms of how students and instructors are assigned to each other. Accounting for non-random variation in which undergraduates enroll in which sections, and who teaches those sections is critical to identifying the true effect.

Coefficients on ACT score and the subject being the undergraduate's major field are also reported in Table 5. Across specifications, higher ACT scores predict more courses taken in any subject, and particularly in the student's major (the interaction of ACT and major). The estimated marginal effect of "In intended major" is zero when ACT score is a very low seven or eight points (such an ACT score is more than three standard deviations below the mean for our sample).

On the right hand side of Table 5 we report estimates of credits taken following exposure to graduate student teaching during only the undergraduate's first term (as opposed to any term). We hypothesize that selection may be weakest during the first undergraduate term since at this early stage new matriculants know the least about differences in subjects, specific courses, and instructor types. Additionally, they may have the least flexibility in scheduling given the need to complete baseline requirements. For this sub-sample, Columns 5-7 replicate Columns 1-3 (there is no Column 4 analog since time cannot vary within a cohort when the analysis is restricted to students' initial fall term). The estimated effect of exposure to graduate student instruction in the first term is, as before, negative.

Selection on Observables

The fixed effects strategies employed in the specifications in Table 5 address selection across departments, courses, and the timing of enrollment in a course. They do not, however, address selection across sections within a course. In Table 6 we explore the extent of this sorting by predicting the likelihood of having

a graduate student instructor using specifications similar to Table 5 except, of course, excluding the graduate student regressor which has become the dependent variable. Students with higher ACT scores are less likely to have a graduate student instructor, though this relationship is practically meaningful only when the subject is the student's intended major. Although the coefficient on "in intended major" is positive and large, undergraduates with ACT scores near or above average do not have a higher probability of seeing a graduate student in their intended major. Additionally, undergraduates attending selective institutions are more likely to have graduate student instructors. These patterns generally hold across fixed effects specifications, and whether considering all courses or focusing just on courses in students' first term.

Instrumental Variables Estimates

Given the sorting concerns observed in Table 6, we remain concerned about issues of selection across course sections as well as other unobserved differences, which our previous OLS specifications fail to adequately address. In response to these lingering sources of potential bias we employ an instrumental variables strategy that uses variation over time in the composition of each department's faculty as an instrument for exposure to graduate student teaching. Ultimately this approach results in estimates quite different from ordinary least squares.

Table 7 reports coefficients and F-tests on excluded instruments from first stages predicting undergraduates' exposure to graduate student instruction as a function of variation in faculty composition within departments over time (Equation 2). As before, the left side of Table 7 uses the sample of all courses while the right side focuses on courses taken during the first term. The pattern across specifications and samples is fairly consistent: the probability of taking a course from a graduate student instructor is significantly related to departmental deviations from the steady-state proportions of most full-time faculty of different ranks,

particularly assistant professors. The results are strongest when focusing on courses during the first term. And, given the coefficient signs, graduate students appear to be a substitute for assistant professor teaching in particular.[11] However, the more conservative standard errors clustered at the section level make the pattern of statistical significance somewhat less consistent. In general, F-tests suggest that the combined excluded instruments do explain substantial variation in undergraduate exposure to graduate student teaching. However, when clustering at the section level, the F-statistics are much lower and sometimes below conventional thresholds.

Table 8 reports two-stage least squares estimates of the effect of graduate student teaching on undergraduates' subsequent interest in the subject. Each cell represents the coefficient of interest from a separate regression. All regressions in Table 8 include campus-by-department-by-course and term fixed effects, as well as the standard set of student-level covariates employed across all analyses. The underlying first stages are reported in Columns 3 and 6 from Table 7 for the left and right sides of Table 8 respectively.

In contrast to the OLS results in Table 5, the 2SLS results suggest that overall—pooling all subjects (Row 1)—taking a first class in from a graduate student increases the probability of majoring in that subject. Across their college career, taking an initial course in a subject from a graduate student increased undergraduates' likelihood of choosing that subject as a major by 9.6 percentage points. This is almost a tripling of the baseline probability of 3.5 percent reported in Column 3 of Table 1. Exposure to graduate student teaching also increased the number of subsequent credits taken in the subject by about 2.3 or about one quarter of a standard deviation. These results are statistically different from zero when we

---

[11] The sometimes positive coefficients for associate professors may reflect the effect of promotions on teaching responsibilities; that is, as faculty are promoted from assistant to associate, their previous course load shifts in part to graduate students.

allow for clustering at the student level, but allowing for clustering at the course-by-section level suggests the estimates may be much noisier. By contrast, it does not appear that undergraduates' major choices or subsequent course taking are influenced when their exposure to graduate instructors occurs during their first college term.

The additional rows of Table 8 report effect estimates for three subsamples: academic subjects (e.g., humanities, sciences, mathematics), and professional subjects (e.g., education, business, engineering). The academic and professional samples are mutually exclusive and exhaustive; the STEM sample draws from both.

The contrast between academic and professional subjects is notable. Undergraduates taught by graduate students are very likely to subsequently major in the subject: an estimated effect of 81 percentage points on major choice. The estimates for course taking behavior are complementary: 83 points more likely to take any additional courses, and an average increase of 30 additional credits in the subject. These differences are robustly significant to both assumptions about error clustering.

Although these results are consistent with the large effects for professional students exposed to adjunct instructors in prior research (Bettinger and Long, 2010), we caution against making strong inferences regarding what appear to be large differences for professional subjects. There is reason to believe that both the timing of undergraduates' first course in a professional subject, as well as the use of graduate student instructors in these departments, differs in substantively important ways compared to academic departments. Turning to focus just on courses in academic disciplines, undergraduates taught by graduates students are about 2.4 percentage points more likely to major in the subject subsequently although estimates for course taking are fairly imprecise.

The results in Table 8 suggest that there are positive impacts on student engagement in subsequent courses. Other research has suggested that high quality

instruction is directly related to subsequent engagement (e.g. Coates 2005). There still remains the possibility that other graduate student characteristics or practices (e.g. ease of grading) could explain the differences in student outcomes noted in Table 8.[12]

Effects of Teaching on Graduate Student Outcomes

While the instrumental variables estimates address issues of selection regarding the impact of graduate student instructors on undergraduates, the estimates in Table 8 are partly a function of the relative quality of graduate student instructors and full-time faculty in our Ohio public institution sample. Of particular concern is that, conditional on hiring graduate student instructors, departments likely do not randomly choose individuals to fill those teaching slots. Faculty may prefer that their most promising students take opportunities to teach as preparation for future careers or to preserve the quality of the department's instruction. Alternatively, faculty may prefer that their most promising students focus on research not teaching. To explore the selection of graduate students into teaching roles, we next estimate the relationship between the frequency with which a graduate student taught courses and the graduate student's own probability of future academic success.

In Column 1 of Table 9 we predict the probability of completing a doctoral degree within six years of starting graduate school. Each additional term in which

---

[12] Appendix Tables 1 and 2 provide two extensions of the analysis in Table 8. First, one potential concern is that faculty assignments vary on intensive margins (for examples, number of students assigned in any one class or in total across sections) in addition to an extensive margin. In Appendix Table 1 we add controls for the class size of student i's focal class, and for the total number of students taught by the instructor teaching student i's focal class during the term. Adding these two additional controls does not change the pattern of results reported in Table 8. Second, Appendix Table 2 reports separate estimates for the effects of "novice" graduate student instructors—those who have not taught previously—and "experienced" graduate student instructors—who have taught at least one prior course. The estimates for these two groups are broadly similar, and we cannot reject the hypothesis that the point estimates are equal.

a graduate student teaches increases the probability of degree completion by 2.1 percentage points in the full sample.[13] The point estimate is the same, 2.1 percentage points, when we limit the sample to graduate students in academic departments, and appears strongest in the humanities. These coefficients suggest that teaching one additional term in the humanities or math is associated with a 13 percent increase in the probability of degree completion within six years.

This predicted increase in degree completion rates could be an effect of the "treatment" of teaching undergraduates, but could just as likely be an artifact of how graduate student instructors are selected. The positive coefficients suggest positive selection: departments may prefer to selectively hire their relatively successful doctoral students to fill teaching slots. Institutions and departments may have heterogeneous preferences and our estimation includes campus-by-department fixed effects in part as a response; however, the results are not much changed without these fixed effects.

Degree completion is a broadly held goal for all graduate students. In addition, many graduate students are also working toward securing employment at a college or university after graduation. Again, opportunities to teach could have a causal effect on graduate students' ability to obtain a post-secondary position. Or alternatively, a positive relationship could suggest that departments select graduate student instructors with preference for those who are on a trajectory toward college or university faculty.

In Column 2 of Table 9 we predict the probability of being employed by an Ohio college or university, other than one's own graduate institution, within six years of starting graduate school. Of note, unlike much of the analysis in this paper,

---

[13] The results in Table 9 were estimated using least squares. We have repeated the same analyses using probit and the estimated marginal effects evaluated at the mean number of terms taught (about two) are very similar in magnitude and precision. Additionally, since our analytic sample is comprised of graduate students who taught undergraduates at least once, it is not appropriate to draw inferences about graduate students who never teach from the results reported in Table 9.

"Ohio" here includes both public and private colleges and universities because ODJFS tracks all employment.[14] Across all departments, we estimate that the probability of subsequently working in higher education increases by 1.4 percentage points for each additional term a graduate student teaches while a candidate; a 9 percent increase over the baseline probability of about 0.15 (see Table 4 Column 1). The estimates for subject-specific sub-samples are mostly not statistically significant. As with the degree completion results, if the mechanism here is selection the selection appears positive, and potentially directed toward students on a future faculty trajectory.

We extend the analysis of future employment by comparing employment at a research university to employment at other colleges and universities.[15] Graduate student teaching is positively correlated with future employment at both research (Column 3) and non-research institutions (Column 4). In the full sample each additional term teaching raises the probability of employment 0.4 percentage points at a research university and 1.1 percentage points at other colleges and universities. If this difference is driven by selection it suggests that departments are more likely to choose graduate students headed toward jobs in colleges and universities where teaching skills are relatively more valuable. If the difference is causal, then the experience of teaching while a graduate student leads more individuals to jobs in college-level teaching.

---

[14] We exclude from the dependent variable employment at the institution where the graduate student earned her degree. The ODJFS records do not distinguish between employment as a graduate student, thus the exclusion is necessary to avoid inadvertently counting student work as obtaining a long-term job. This exclusion may introduce some measurement error to the extent a university is willing to hire its own graduates as faculty. Additionally, the ODJFS records do not distinguish job roles, thus our dependent variable includes both faculty and administrative employment at a college or university.

[15] We separated institutions using the 2005 Carnegie classifications. For "research university" we include RU/VH, Research University / Very High Research Activity, as well as RU/H, Research University / High Research Activity. All other Ohio post-secondary institutions are included in "non-research college or university."

The mechanisms behind the results in Table 9 could include some causal role for teaching opportunities. For example, graduate students who teach more frequently could develop skills preferred by non-research-focused institutions, or could miss out on opportunities to develop research skills needed for securing employment at research universities. These effects, if they exist at all, are relative since the results in columns 3 and 4 of Table 9 suggest that teaching predicts a greater chance of employment at both types of institutions.

Which interpretation—selection or causation—is correct? For our purposes, it may not matter. Both interpretations give the result that graduate student teaching generates positive results. If the results represent causal effects, then graduate student teaching is a positive experience and helps graduate students graduate in a timely way and gain employment in higher education, at least in their initial careers. Graduate student teaching leads to positive effects on graduate students and mostly positive or no effects on the undergraduate students they teach. If, on the other hand, the results represent selection of university-bound graduate students, then departments are successfully identifying the graduate students who will have positive effects on students, or at least no differential effects compared to full-time faculty.

V. Conclusion

This paper makes two important contributions to evidence on the effects of graduate student teaching. First, we compare outcomes for undergraduates taught by graduate students to outcomes for their peers taught by full-time faculty. We find that undergraduates who take their first course in a given subject from a graduate student are nearly twice as likely to subsequently major in that subject compared to their peers who take the same course from full-time faculty. Differences in major choice do not arise, however, when the course is taken during the undergraduate's first term in college. Similarly, in our analysis setting, there are

no statistically significant differences in the subsequent course taking decisions of undergraduates taught by full-time faculty or graduate students when pooling across subjects, although there is some heterogeneity from subject to subject in these estimated effects. Prior research suggests that subsequent engagement is an indicator of teaching quality although other explanations are possible.

These generally null or positive estimates differ noticeably from simpler comparisons which show negative effects on undergraduates even when comparing different sections of the same course in the same term. Our estimates use an instrumental variables approach to address the likely substantial non-random sorting of undergraduates and instructors across courses and sections. And our results reinforce the first-order concern of such sorting when studying the effects of teachers on students.

Second, we compare outcomes for graduate students themselves who teach more or less frequently. We find that graduate students who teach more frequently are more likely to graduate in a timely manner and more likely to subsequently be employed by a college or university in their early careers. While we are less certain whether these results arise from selection or some causal mechanism, both explanations are consistent with an optimistic view of status quo graduate student teaching. We do not, however, have measures of longer-run outcomes or of teaching or research productivity in these new careers.

These two sets of results—outcomes for undergraduates and outcomes for graduate students—are important for researchers and policy makers to consider simultaneously. Universities face a set of interrelated production functions; inferences or prescriptions that ignore the production of graduate students when considering the state of undergraduate education are taking a short-run perspective. But the tradeoffs between undergraduate and graduate education are not necessarily severe. Our results suggest, at least for systems similar to Ohio's, that universities can find an equilibrium which is productive on both fronts.

References

Bair, Carolyn R. and Jennifer G. Haworth. 2004. "Doctoral Student Attrition and Persistence: A Meta–Synthesis of Research." In *Higher Education: Handbook of Theory and Research (Vol. XIX)*, J.C. Smart (Ed.). New York: Agathon Press.

Bettinger, Eric P. 2004. "Is the Finish Line in Sight? Financial Aid's Impact on Retention and Graduation" in Caroline M. Hoxby (Ed.), *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It.* Chicago: University of Chicago Press and NBER.

Bettinger, Eric P. and Bridget Terry Long. 2010. "Does Cheaper Mean Better? The Impact of Using Adjunct Instructors on Student Outcomes." *Review of Economics and Statistics,* 92(3):598-613.

Bettinger, Eric P. and Bridget Terry Long. 2006. "The Increasing Use of Adjunct Instructors at Public Institutions: Are We Hurting Students?" in *What's Happening top Public Higher Education?* Ronald G. Ehrenberg (Ed.). Westport, CT: American Council on Education / Praeger.

Borjas, George J. 2000. "Foreign-Born Teaching Assistants and the Academic Performance of Undergraduates." *American Economic Review* 90(2):355-359.

Bowen, William G. and Neil L. Rudenstine. 1992. *In Pursuit of the PhD.* Princeton, N.J.: Princeton University Press.

Carrell, Scott E. and James E. West. 2010. "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors." *Journal of Political Economy* 118(3):409-432.

Coates, Hamish. "The value of student engagement for higher education quality assurance." Quality in Higher Education 11.1 (2005): 25-36.

Ehrenberg, Ronald G., and Panagiotis G. Mavros. 1995. "Do Doctoral Students' Financial Support Patterns Affect Their Times-To-Degree and Completion Probabilities?" *Journal of Human Resources* 30(3):581-609.

Fleisher, Belton, Masanori Hashimoto, and Bruce A. Weinberg. 2002. "Foreign GTAs Can Be Effective Teachers of Economics." *Journal of Economic Education* 33(4):299-325.

Hanushek, Eric A. and Steven G. Rivkin. 2010. "Generalizations About Using Value-added Measures of Teacher Quality." *American Economic Review* 100(2):267-271.

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

Jacobs, Lucy C. and Charles B. Friedman. 1988. "Student Achievement Under Foreign Teaching Associates Compared with Native Teaching Associates." *The Journal of Higher Education* 59(5):551-563.

Knotts, H. Gibbs and Eleanor C. Main. 1999. "Teaching Ph.D. Students to Teach Political Science: The Emory TATTO Program." *PS: Political Science and Politics* 32(3):607-610.

Koehnecke, Dianne Swenson. 1991. "Boundaries of Graduate Assistants: Bouncing off Boundaries." Paper presented at the Annual Meeting of the Conference on College Composition and Communication, Boston, MA.

Mattson, Kevin. 2000. "The Academic Labor Movement: Understanding its Origin and Current Challenges. *Social Policy.* 30(4):4-10.

McLeod, Susan H. and Fred S. Schwarzbach. 1993. "What about the TAs? Making the Wyoming Resolution a Reality for Graduate Students." *WPA: Writing Program Administration* 17(1-2):83-87.

Meyers, Steven A. and Loreto Prieto R. 2000. "Training in the Teaching of Psychology: What is Done and Examining the Differences." *Teaching of Psychology.* 27(4):258-261.

Mortenson, Thomas. 2002. "Chance for College by Age 19 by State in 2000." Oskaloosa, IA: Postsecondary Education Opportunity.

Norris, Timothy. 1991. "Nonnative English-speaking Teaching Assistants and Student Performance." *Research in Higher Education* 32(4).

Prieto, Loreto R. and Elizabeth M. Altmaier. 1994. "The Relationship of Prior Training and Previous Teaching Experience to Self-Efficacy among Graduate Teaching Assistants." *Research in Higher Education* 35(4):481-97.

Reeve, Johnmarshall, and Hyungshim Jang. "What teachers say and do to support students' autonomy during a learning activity." *Journal of educational psychology* 98.1 (2006): 209.

Seagram, Belinda C., Judy Gould, and Sandra W. Pyke. 1998. "An Investigation of Gender and Other Variables on Time to Completion of Doctoral Degrees." *Research in Higher Education* 39 (3): 319-335.

Sharnoff, Elena. 1993. "Neither Fish Nor Fowl: Graduate Students, Unionization, and the Academy" Talk presented at the Annual Meeting of the Modern Language Association, Toronto, Canada, Dec 1993.

Skinner, Ellen A., and Michael J. Belmont. "Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year." *Journal of educational psychology* 85.4 (1993): 571.

Slevin, James F. 1992. *The Next Generation: Preparing Graduate Students for the Professional Responsibilities of College Teachers.* Washington, D.C.:Association of American Colleges.

Smith, Kathleen S. 2001. "Pivotal Events in Graduate Teacher Preparation for a Faculty Career." *Journal of Graduate Teaching Assistant Development* 8(3):97-105.

Vaughn, Willliam (1998) "Apprentice or Employee? Graduate Students and Their Unions." *Academe* 84(6):43-49.

Watts, Michael, and Gerald J. Lynch. 1989. "The Principles Course Revisited." *American Economic Review* 79(2):236-41.

Wise, Max. 1967. "Who Teaches the Teachers?" in *Improving College Teaching*. Washington, D.C.: American Council on Education.

Table 1: Full-time, Traditional-aged Students at Four-year, Public Colleges in Ohio

	Sample			
			Courses with Graduate Instructor Variation	
	All Initial Courses	Initial Courses During First Term	All Initial Courses	Initial Courses During First Term
	(1)	(2)	(3)	(4)
A. Student Characteristics				
Male	0.4499	0.4588	0.4427	0.4425
White	0.8589	0.8605	0.8623	0.8700
Black	0.0731	0.0725	0.0704	0.0682
ACT composite score (36 max)	22.7337 (4.1095)	22.7207 (4.1045)	23.0634 (4.0094)	22.7306 (3.9335)
Ohio resident	0.9986	0.9985	0.9987	0.9985
Entering fall 1999 cohort	0.4191	0.4103	0.4571	0.4353
Selective institution	0.6291	0.5774	0.7282	0.6440
Student Observations	43,177	42,742	39,920	35,718
B. Student-by-subject Characteristics				
Initial Course in Subject				
Class taken in intended major	0.0334	0.0494	0.0226	0.0329
Number of credits taken in the subject that term	3.0676 (1.0118)	3.0846 (1.0069)	3.1479 (0.9065)	3.1512 (0.9320)
Adjunct instructor	0.2266	0.2567	0.2278	0.2906
Graduate student instructor	0.2103	0.2110	0.4360	0.3775
Subsequent Behavior in Subject				
Any additional credits	0.4889	0.6736	0.5419	0.7213
Subsequent credits taken within the next year	4.8575 (3.682)	5.8935 (3.7254)	5.0836 (3.6216)	6.0738 (3.5479)
Total subsequent credits	4.4773 (8.5305)	7.0652 (10.4902)	4.8390 (8.4511)	7.2017 (10.0204)
Majored in subject within four yrs	0.0517	0.0783	0.0352	0.0529
Student-by-Subject Observations	409,308	138,920	176,360	70,245

Note: Standard deviations are shown in the parentheses. Sample is restricted to first-time, full-time freshmen who are of traditional age (18–20 years old) and entered a public, four-year college in Ohio during fall 1998 or fall 1999. Students must also have taken the ACT. Selective institutions are defined as "competitive" institutions by Barron's Educational Guides (1997) and include Bowling Green State University, Miami University, Ohio State University, Ohio University, and Youngstown State University. The sample of courses is limited to the first courses taken in a subject and, in the right hand columns, students who took courses with at least 10% and less than 90% of the sections taught by graduate student instructors.

Table 2: Proportion of Students with a Graduate Student Instructor by Subject

	Full Sample	Courses with Graduate Instructor Variation
	(1)	(2)
All Departments	0.2103	0.4360
Humanities	0.3139	0.4684
English	0.4001	0.4447
History	0.2435	0.4938
Foreign Language	0.2965	0.4792
Social Sciences	0.2095	0.4689
Economics	0.1119	0.3371
Political Science	0.1793	0.4649
Psychology	0.2995	0.5221
Sociology	0.1965	0.3666
Sciences	0.0843	0.2646
Biology	0.1019	0.2240
Chemistry	0.1051	0.2659
Physics	0.1018	0.3139
Math and Statistics	0.2174	0.3467
Engineering	0.0631	0.2113
Computer Science	0.1845	0.5304
Communications and Journalism	0.0717	0.3323
Business	0.0293	0.2545
Education	0.3112	0.5409
Social Work and Public Affairs	0.0692	0.3258

Notes: The subgroups shown under the major school groupings are not a complete list of departments. The sample of courses is limited to the first courses taken in a subject and, in the right hand column, students who took courses with at least 10% and less than 90% of the sections taught by graduate student instructors.

Table 3: Faculty Characteristics

	All Instructors		Instructors Teaching a First Course in a Subject		
	Raw Means	Weighted by Course Enrollment	Full-time Instructors	Adjunct Instructors	Graduate Students
	(1)	(2)	(3)	(4)	(5)
Percent of instructors					
with a Ph.D.	24.1	44.1	64.6	22.3	3.5
working part time	33.4	21.1	--	100.0	--
graduate students	18.1	14.1	--	--	100.0
non-tenure track	60.6	49.0	60.1	97.8	91.1
female	38.4	35.6	31.6	47.2	53.3
minority	21.8 [41,685]	18.2 [2,159,524]	18.5 [7,422]	19.9 [5,380]	36.1 [6,326]
Year Born	1957.2 [37,419]	1953.8 [1,953,370]	1952.7 [6,645]	1955.3 [4,770]	1969.0 [5,847]
Observations	49,343	2,366,861	7,947	5,449	6,378

Note: Sample restricted to active faculty teaching between 1998 and 2003 at the undergraduate level. Statistics calculated using less than the full sample have the observations in brackets.

Table 4: Graduate Student Characteristics

	Full Sample	Academic Departments
	(1)	(2)
Terms taught within six years		
One	0.2201	0.1886
Two	0.2213	0.2256
Three	0.1858	0.1865
Four	0.1101	0.1082
Five or more	0.2627	0.2911
Received PhD within six years	0.1638	0.1751
Employed within six years		
Any Ohio college or university	0.1500	0.1409
Research university	0.0504	0.0505
Non-research college or university	0.1134	0.1032
Any sector in Ohio	0.5257	0.4947
Enrolled during sixth year	0.2989	0.3523
Number of graduate students	2,680	1,405

Note: Sample restricted to graduate students who began their graduate program in 1998 or 1999, and subsequently taught at least one undergraduate class within six years.

Table 5: OLS Estimates of the relative Effect of Having a Graduate Student Instructor
Dependent Variable: Total Subsequent Credit Hours

| | Total subsequent credit hours in subject | | | | | | |
| | Courses Taken Any Term | | | | Initial Term Only | | |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Graduate student instructor	-0.575**	-0.213**	-0.116**	-0.263**	-0.334**	-0.275**	-0.414**
	(0.047)	(0.044)	(0.044)	(0.053)	(0.090)	(0.088)	(0.094)
	[0.130]	[0.072]	[0.057]	[0.068]	[0.245]	[0.145]	[0.126]
ACT score	0.077**	0.069**	0.032**	0.029**	0.100**	0.070**	0.038**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.010)	(0.010)	(0.010)
	[0.008]	[0.006]	[0.006]	[0.006]	[0.017]	[0.013]	[0.012]
In intended major	-6.103**	-2.521	-8.275**	-7.991**	-8.102**	-1.203	-10.376**
	(1.753)	(1.653)	(1.639)	(1.677)	(2.537)	(2.421)	(2.466)
	[2.581]	[1.735]	[1.719]	[1.791]	[2.976]	[2.374]	[2.400]
In intended major * ACT score	0.858**	0.545**	0.642**	0.598**	0.994**	0.535**	0.767**
	(0.075)	(0.071)	(0.071)	(0.072)	(0.108)	(0.104)	(0.106)
	[0.128]	[0.073]	[0.072]	[0.075]	[0.133]	[0.103]	[0.100]
Fixed effects							
Campus		Y				Y	
Department		Y				Y	
Campus-by-department-by-course			Y				Y
Year-term		Y	Y				
Campus-by-dept-by-course-by-term				Y			
Observations	176,360	176,360	176,360	176,360	70,245	70,245	70,245
R-squared	0.086	0.181	0.041	0.023	0.090	0.148	0.028

Note: Standard errors that cluster at the student level due to the multiple observations per student are shown in the parentheses. However, because the treatment (instructor type) varies by section, we also report standard errors that cluster at the section level in brackets. The models include the following covariates: race dummy variables, gender dummy variable, a dummy variable for being an Ohio resident, the total number of credits taken in the subject that term, a dummy variable for the fall 1999 cohort, and a dummy variable for whether the institution is selective.
** Significant at the 1 percent level   * Significant at the 5 percent level   + Significant at the 10 percent level

Table 6: OLS Estimates of the Probability of Having a Graduate Student Instructor in a First Course in a Subject

	Courses Taken Any Term				Initial Term Only		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ACT score	-0.001**	-0.003**	-0.003**	-0.003**	-0.004**	-0.002**	-0.003**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
	[0.001]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]	[0.001]
Course in intended major	0.342**	0.214**	0.092*	0.057+	0.388**	0.262**	0.147**
	(0.040)	(0.039)	(0.038)	(0.034)	(0.052)	(0.052)	(0.050)
	[0.066]	[0.053]	[0.048]	[0.043]	[0.089]	[0.072]	[0.062]
Course in intended major * ACT	-0.018**	-0.010**	-0.004**	-0.003*	-0.018**	-0.012**	-0.008**
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
	[0.003]	[0.002]	[0.002]	[0.002]	[0.004]	[0.003]	[0.003]
White student	0.002	0.005	0.002	0.005	0.031**	0.022**	0.016*
	(0.005)	(0.005)	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)
	[0.005]	[0.005]	[0.004]	[0.004]	[0.008]	[0.008]	[0.007]
Black student	0.007	-0.008	-0.004	-0.007	0.009	0.010	0.004
	(0.007)	(0.006)	(0.006)	(0.005)	(0.010)	(0.010)	(0.009)
	[0.007]	[0.007]	[0.006]	[0.006]	[0.013]	[0.012]	[0.010]
Male student	0.000	0.004+	0.006**	0.005*	-0.006	0.007+	0.008*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.003)
	[0.003]	[0.003]	[0.002]	[0.002]	[0.006]	[0.005]	[0.004]
In state student	0.040	-0.004	-0.011	0.005	0.024	-0.000	-0.028
	(0.034)	(0.031)	(0.030)	(0.028)	(0.049)	(0.044)	(0.041)
	[0.031]	[0.030]	[0.030]	[0.026]	[0.045]	[0.043]	[0.041]
Number of credits taken that term in the subject	-0.027**	-0.005**	-0.025**	-0.025**	-0.013**	0.003	-0.025**
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
	[0.004]	[0.004]	[0.003]	[0.003]	[0.006]	[0.007]	[0.005]
1999 cohort	0.047**	-0.007*	-0.020**	-0.014**	0.061**	0.060**	0.066**
	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
	[0.010]	[0.009]	[0.008]	[0.007]	[0.018]	[0.018]	[0.020]
Selective institution	0.146**				0.068**		
	(0.003)				(0.004)		
	[0.013]				[0.020]		

Fixed effects							
Campus		Y				Y	
Department		Y				Y	
Campus-by-department-by-course			Y				Y
Year-term		Y	Y				
Campus-by-dept-by-course-by-term				Y			
Observations	176,360	176,360	176,360	176,360	70,245	70,245	70,245

Note: Standard errors that cluster at the student level are shown in the parentheses. Standard errors that cluster at the section level are shown in brackets. The models also include a dummy variable for being an Ohio resident.

** Significant at the 1 percent level   * Significant at the 5 percent level   + Significant at the 10 percent level

Table 7: Results of the First-Stage IV Regressions: Probability of Having a Graduate Student Instructor

	Courses Taken Any Term			Initial Term Only		
	(1)	(2)	(3)	(4)	(5)	(6)
Proportion assistant professors	-0.098**	-0.399**	-0.399**	-0.615**	-0.814**	-0.958**
	(0.010)	(0.012)	(0.045)	(0.023)	(0.028)	(0.054)
	[0.041]	[0.045]	[0.055]	[0.115]	[0.165]	[0.359]
Proportion associate professors	0.187**	-0.061**	-0.061	-0.091**	0.011	0.067
	(0.009)	(0.011)	(0.040)	(0.020)	(0.027)	(0.047)
	[0.038]	[0.040]	[0.045]	[0.098]	[0.122]	[0.200]
Proportion full professors	-0.022**	-0.070**	-0.070**	-0.185**	-0.095**	-0.577**
	(0.005)	(0.007)	(0.024)	(0.010)	(0.016)	(0.030)
	[0.019]	[0.024]	[0.027]	[0.038]	[0.075]	[0.105]
Fixed effects						
Campus		Y			Y	
Department		Y			Y	
Campus-by-dept-by-course			Y			Y
Year-term		Y	Y		Y	Y
Observations	176,360	176,360	176,360	70,245	70,245	70,245
R-squared	0.025	0.095	0.015	0.027	0.109	0.027
F-test on excluded instruments (clustered at student level)	173.853	360.903	137.509	410.956	285.419	200.247
F-test on excluded instruments (clustered at section level)	9.066	26.020	9.653	15.616	9.771	11.241

Note: Standard errors that cluster at the student level are shown in the parentheses. Standard errors that cluster at the section level are shown in brackets. The models include the following covariates: ACT composite score, whether the course is in the student's intended major, the interaction between ACT and intended major, race dummy variables, gender dummy variable, a dummy variable for being an Ohio resident, the total number of credits taken in the subject that term, a dummy variable for the fall 1999 cohort, and a dummy variable for whether the institution is selective (this drops out when using fixed effects). Not reported is the analogous first stage for exposure to an adjunct instructor.
** Significant at the 1 percent level   * Significant at the 5 percent level   + Significant at the 10 percent level

Table 8: IV Estimates of the Effects of Graduate Student Instructors on Subsequent Interest

	Courses taken any term			Initial term only		
	Major choice	Any additional courses in subject	Total subsequent credits	Major choice	Any additional courses in subject	Total subsequent credits
	(1)	(2)	(3)	(4)	(5)	(6)
All subjects	0.096** (0.026) [0.070]	0.025 (0.062) [0.148]	2.305* (0.969) [2.425]	0.014 (0.023) [0.031]	-0.053 (0.039) [0.047]	-0.846 (1.037) [1.134]
Academic subjects	0.024+ (0.014) [0.032]	-0.037 (0.051) [0.102]	-0.289 (0.615) [0.990]	0.016 (0.021) [0.025]	-0.053 (0.037) [0.045]	-0.097 (0.901) [1.022]
Professional subjects	0.806** (0.120) [0.276]	0.832** (0.141) [0.281]	30.047** (4.740) [10.800]	1.320** (0.453) [0.836]	0.211 (0.293) [0.287]	11.952 (14.387) [16.452]

Note: Each cell reports the coefficient of interest from a separate regression. Standard errors that cluster at the student level due to the multiple observations we have per student are shown in the parentheses. However, because the treatment (instructor type) varies by section, we also report standard errors that cluster at the section level in brackets. The models include campus-by-department-by-course fixed effects, and term fixed effects. They also include the following covariates: ACT composite score, whether the course is in the student's intended major, the interaction between ACT and intended major, race dummy variables, gender dummy variable, a dummy variable for being an Ohio resident, the total number of credits taken in the subject that term, and a dummy variable for the fall 1999 cohort.
** Significant at the 1 percent level   * Significant at the 5 percent level   + Significant at the 10 percent level

Table 9: The Effect of Teaching an Additional Term on the Predicted Probability
of Graduate Students' Own Academic Success within Six Years

| | Received Doctoral Degree | Employed by a College or University in Ohio | | | Sample Size |
		Any College or University	Research University[a]	Non-Research College or University[a]	
	(1)	(2)	(3)	(4)	(5)
Any Department	0.021** (0.005)	0.013** (0.003)	0.004* (0.002)	0.011** (0.003)	2,680
Academic Departments	0.021** (0.006)	0.008+ (0.005)	0.002 (0.003)	0.007+ (0.004)	1,405
Humanities	0.037** (0.008)	0.013 (0.008)	0.003 (0.005)	0.010 (0.007)	588
Social Sciences	0.008 (0.009)	0.011 (0.010)	-0.002 (0.002)	0.016 (0.010)	295
Sciences	-0.002 (0.013)	0.000 (0.008)	0.006 (0.005)	-0.001 (0.008)	366
Math	0.020 (0.015)	-0.000 (0.011)	0.003 (0.009)	-0.003 (0.007)	186

Note: Each cell represents the coefficient on "Number of terms taught" during the graduate student's first six years after beginning graduate school. Each is from a separate campus-by-department fixed effects specification. Clustered (campus-by-department) standard errors in parentheses.
[a] "Research University" includes institutions with a 2005 Carnegie classification of RU/VH, Research University / Very High Research Activity, as well as RU/H, Research University / High Research Activity. All other Ohio post-secondary institutions are included in "non-research college or university."
** Significant at the 1 percent level   * Significant at the 5 percent level   + Significant at the 10 percent level

Appendix (to be posted online)

Appendix Table 1: IV Estimates of the Effects of Graduate Student Instructors on Subsequent Interest, Controlling for Class Size and Professor Teaching Load

	Courses taken any term			Initial term only		
	Major choice	Any additional courses in subject	Total subsequent credits	Major choice	Any additional courses in subject	Total subsequent credits
	(1)	(2)	(3)	(4)	(5)	(6)
All subjects	0.081**	0.025	1.794*	0.029	-0.030	-1.308
	(0.021)	(0.048)	(0.816)	(0.034)	(0.055)	(1.686)
	[0.055]	[0.086]	[1.820]	[0.047]	[0.058]	[1.830]
Academic subjects	0.015	0.011	-0.173	0.049+	-0.029	0.320
	(0.013)	(0.043)	(0.576)	(0.028)	(0.049)	(1.294)
	[0.028]	[0.068]	[0.886]	[0.040]	[0.052]	[1.462]
Professional subjects	0.635**	0.456**	22.778**	2.023+	0.470	23.640
	(0.092)	(0.100)	(3.730)	(1.076)	(0.536)	(26.019)
	[0.195]	[0.155]	[7.788]	[2.074]	[0.602]	[41.485]

Note: Each cell reports the coefficient of interest from a separate regression. Standard errors that cluster at the student level due to the multiple observations we have per student are shown in the parentheses. However, because the treatment (instructor type) varies by section, we also report standard errors that cluster at the section level in brackets. Estimation identical to Table 8, except each regression also includes controls for class size (the number of students in student i's first class) and the professor's student load (the total number of students taught by student i's first class professor j in the same term of student i's first class).
** Significant at the 1 percent level   * Significant at the 5 percent level   + Significant at the 10 percent level

Appendix Table 2: IV Estimates of the Effects of Graduate Student Instructors on Subsequent Interest, Separate Estimates for Graduate Student Teaching Experience

	Courses taken any term			Initial term only		
	Major choice	Any additional courses in subject	Total subsequent credits	Major choice	Any additional courses in subject	Total subsequent credits
	(1)	(2)	(3)	(4)	(5)	(6)
All subjects						
Graduate instructor	0.098*	-0.075	3.178+	-0.040	0.163	-5.750
no prior experience	(0.047)	(0.135)	(1.645)	(0.079)	(0.183)	(4.613)
	[0.124]	[0.344]	[3.844]	[0.106]	[0.283]	[7.197]
Graduate instructor	0.095+	0.160	1.194	0.032	-0.126	0.793
some experience	(0.055)	(0.158)	(1.966)	(0.037)	(0.079)	(1.731)
	[0.160]	[0.453]	[5.100]	[0.049]	[0.121]	[3.084]
Academic subjects						
Graduate instructor	0.032	-0.145	1.007	0.127*	0.146	0.802
no prior experience	(0.026)	(0.104)	(1.078)	(0.063)	(0.136)	(2.656)
	[0.054]	[0.233]	[1.602]	[0.142]	[0.216]	[3.035]
Graduate instructor	0.017	0.069	-1.516	-0.034	-0.145+	-0.501
some experience	(0.025)	(0.099)	(1.068)	(0.039)	(0.077)	(1.470)
	[0.056]	[0.250]	[1.912]	[0.084]	[0.121]	[1.855]
Professional subjects						
Graduate instructor	1.029**	1.034**	34.083**	1.165*	0.125	5.382
no prior experience	(0.371)	(0.397)	(12.480)	(0.573)	(0.316)	(17.440)
	[0.843]	[0.885]	[28.945]	[1.239]	[0.490]	[28.486]
Graduate instructor	1.368	1.344	40.187	2.116*	0.655	45.768
some experience	(0.857)	(0.916)	(28.988)	(1.016)	(0.689)	(30.990)
	[1.853]	[1.945]	[64.044]	[2.023]	[0.892]	[50.888]

Note: Each cell reports the coefficient of interest from a separate regression. Standard errors that cluster at the student level due to the multiple observations we have per student are shown in the parentheses. However, because the treatment (instructor type) varies by section, we also report standard errors that cluster at the section level in brackets. Estimation identical to Table 8, except that the endogenous indicator of interest "taught by graduate student" is separated into two indicators "taught by a graduate student with no prior teaching experience" and "taught by a graduate student who has taught in at least one prior term."
** Significant at the 1 percent level   * Significant at the 5 percent level   + Significant at the 10 percent level