

## Article

# The Student Evaluation of Teaching Premium for Clinical Faculty in Economics

Jasmine Bordere <sup>1</sup>, Fonda Carter <sup>1</sup>, Steven Caudill <sup>2</sup> and Franklin Mixon, Jr. <sup>1,\*</sup> 

<sup>1</sup> Department of Accounting and Finance, Columbus State University, Columbus, GA 31907, USA; bordere\_jasmine@columbusstate.edu (J.B.); carter\_fonda@columbusstate.edu (F.C.)

<sup>2</sup> Department of Economics, Florida Atlantic University, Boca Raton, FL 33431, USA; scaudill@fau.edu

\* Correspondence: mixon\_franklin@columbusstate.edu

**Abstract:** This article uses student evaluation of teaching (SET) data for 947 faculty members affiliated with 90 U.S. colleges and universities to study the presence of a teaching quality rating premium for clinical economics faculty relative to traditional tenure-track economics faculty. Based on OLS estimation, we find this difference ranges between 3.9% and 4.8% and is robust to different econometric model specifications. Moreover, the average treatment effect from a propensity score weighting approach suggests that the difference ranges between 5.8% and 6.1%. Lastly, our analysis produces an institutional ranking of economics departments based on department-level SETs. Overall, our findings are encouraging signs for the hiring and retention of clinical faculty in economics departments.

**Keywords:** economics education; teaching quality; clinical faculty; student evaluations of teaching

## 1. Introduction

Clinical academic positions and other non-tenure-track positions are becoming more prevalent in U.S. economics departments. Pieters and Roark [1] find that the average economics department currently employs almost the same number of tenure-track assistant professors as non-tenure-track professors. The teaching focus and contingent character of those positions are clearly stated in their job descriptions [2]. Asali [3] argues that non-tenure-track appointments complement tenure-track professorships by allowing the latter to concentrate on the increasingly more challenging work of academic research. Whether or not traditional tenure-track faculty reap the intended publication benefits of this labor management strategy is up for debate [4]. Evidence of improved student outcomes derived from receiving instruction from non-tenure-track faculty is also inconclusive. Ran and Xu [5] find that contingent faculty have positive impacts on current course grades but negative impacts on subsequent course outcomes.

The comprehensive analysis of public Ph.D.-granting economics departments conducted by Hilmer and Hilmer [6] found these non-tenure-track instructors tend to be younger, are more likely to be female and are more likely to teach in the program from which they received their Ph.D. when compared to their traditional tenure-track colleagues. Moreover, they were assigned to teach both more courses and many more students. Survey evidence collected in other fields describes non-tenure-track positions as “the lowest academic rank, with short contracts and unclear expectations” [7]. These might not be the best conditions to develop the expected high-quality teaching practices associated with those positions.

Student evaluations are a standard component of the way colleges and universities assess the quality of an instructor’s teaching for purposes of promotion and tenure, merit raise allocations, and reappointment. That motivates our research question: how do clinical faculty in economics stack up against traditional economics faculty in terms of student evaluations of teaching? With an eye to understanding how student evaluations of clinical



**Citation:** Bordere, J.; Carter, F.; Caudill, S.; Mixon, F., Jr. The Student Evaluation of Teaching Premium for Clinical Faculty in Economics. *Educ. Sci.* **2024**, *14*, 107. <https://doi.org/10.3390/educsci14010107>

Academic Editor: Xinqiao Liu

Received: 20 December 2023

Revised: 11 January 2024

Accepted: 12 January 2024

Published: 18 January 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

faculty in economics compare to those for traditional tenure-track economics faculty, we next turn our attention to the economic education literature on the determinants of student evaluation scores.

## 2. Prior Literature

Student evaluations of teaching (SETs) reflect a broad range of objective and subjective qualities of instructors. These relate to academic discipline, gender and other demographics, and teaching experience. Recent economic education research pertaining to each of these qualities is reviewed in the sub-sections below. The final two sub-sections below discuss the efficacy of the use of SET data available from RateMyProfessors.com (RMP) in academic research and how those data have been employed in prior studies.

### 2.1. Academic Discipline Effects in SETs

Ongeri [8] takes a deep dive into the SET literature published over the prior 20 years, which consistently puts the quality of instruction in economics below that in most other academic disciplines. As indicated by the review, the relatively lower quality of instruction in economics tends to be the result of its math orientation, the lack of organization and presentation skills of its instructors, the utilization of multiple-choice tests in assessing economics literacy, and employment of “chalk and talk” teaching approaches that require little, if any, active participation by students. As Asarta et al. [9] point out, this intensive use of lecturing is a sustained practice in the economics discipline. Ongeri [8] concludes that low SETs in economics courses likely reflect a real, underlying problem with the adequacy of university economics education. Relatedly, econometric analysis of a panel of SET data from undergraduate economics courses presented in other studies finds that SETs are a function of class size [10–12] and instructor age [11]. Not only does rapport between instructors and students deteriorate with increasing class size, but large lecture enrollments require the use of the types of standardized tests that Ongeri [8] explains reduce SET scores. Additionally, the finding in McPherson et al. [11] that economics students prefer younger instructors to older instructors is consistent with the finding in Ongeri [8] of the low regard held by students of antiquated teaching methodologies.

### 2.2. Gender and Other Demographic Effects in SETs

The potential for gender discrimination in SETs has been the subject of a number of studies. The study of a potential gender effect in SETs by Wagner et al. [13] is perhaps the most comprehensive of its type. It analyzes a unique dataset featuring mixed teaching teams and a diverse group of students and teachers. The blended co-teaching approach allows for the examination of the link between SETs and instructor gender (and ethnicity) in a way that encompasses *within*-course variations. The analysis finds a negative effect of being a female instructor on SETs equal to about 25% of the sample standard deviation of SETs [13]. More specifically, the results suggest that female instructors are 11 percentage points less likely to meet the SET threshold for promotion to associate professor than are their male counterparts [13]. Boring’s [14] application of SET data from a French university to both fixed effects and generalized ordered logit regression analyses finds that male students express a bias in favor of male professors. Among the individual teaching dimensions, Boring [14] finds that students’ SETs match gender stereotypes. For example, male professors are perceived by both male and female students as being more knowledgeable in the subject and exhibiting superior class management. Interestingly, the analyses also suggest that students appear to learn as much from female professors as they do from male professors [14].

A recent study by Mengel et al. [15] employs a quasi-experimental dataset of 19,952 student evaluations of university faculty in a context where students are randomly allocated to female or male instructors. Even though students’ grades and effort are unaffected by the instructor’s gender, the results suggest that female instructors receive systematically lower SETs than their male colleagues. This bias is driven by male students’ evaluations

and is particularly pronounced for junior female faculty [15]. Mengel et al. [15] add that gender bias in teaching evaluations may alter the career progression of women by affecting junior women's confidence, as well as through the reallocation of instructor resources away from research and toward teaching. Keng [16] examines SET data from a public university in Taiwan in order to test for statistical discrimination against female instructors. In doing so, the study relies upon a learning model wherein the instructors' value added to grades is used to measure teaching effectiveness. Empirical results support the presence of gender bias in SETs, particularly by male students and in science- and math-oriented academic disciplines that employ relatively few female instructors [16]. As one moves toward academic disciplines wherein female instructors are more common, the ratings of female instructors by female students rise. Lastly, in light of earlier findings of a gender effect in SETs, Buser et al. [17] investigate whether gender differences in SETs vary over the course of a semester. Their study examines the application of SETs in principles of economics courses at multiple institutions on three separate occasions during the semester in order to determine whether the evaluations of male and female instructors change throughout an academic term, specifically after the first exam is returned. Tests presented by Buser et al. [17] point toward a negative effect on evaluations for female instructors relative to male instructors associated with returning grades, thus highlighting the importance of temporal effects (related to gender) in the application of SETs.

A number of studies have found that race/ethnicity bias is also present in SETs (e.g., [18–22]). One example is the experimental approach employed by Chisadza et al. [18] that randomly assigned South African students to various course lecturers who all used the same instructional materials. They report that black lecturers received lower SETs than white lecturers, and that this result held even for black students [18]. Relatedly, Chávez and Mitchell [20] utilize a quasi-experimental design wherein instructors recorded welcome videos that were presented to students at the beginning of an online course, thus revealing instructors' race/ethnicity. Examination of post-course SETs revealed that non-white instructors received lower SETs than their white counterparts, holding constant course content, assignments, schedules and communications [20]. A similar study by Basow et al. [23] indicates that brief lectures presented by computer-animated instructors who vary by race are associated with biased SETs, where the white animated actors receive higher SETs than their black counterparts.

Lastly, a recent strand of the literature focuses on the abusive nature of SETs, particularly with regard to anonymized open-ended comments from students (e.g., [24]). Extensive studies by Jones et al. [25], Tucker [26] and Uttl and Smibert [27] document that abusive comments are often present in SETs, and that most are directed towards women and other minority groups. New research by Heffernan [24] examines results from a survey of 674 academics about abusive comments as well as the anonymized student comments attached to SETs at the 16,000 higher education institutions that collect this information each semester. According to the survey, 59% of academics report having been the target of abusive language in open-ended comments attached to SETs. As a result, two-thirds of this group reported mental health declines, while about one-sixth of this group sought professional medical help [24]. Heffernan's [24] results support prior work by Jones et al. [25], Tucker [26] and Uttl and Smibert [27] by indicating that the brunt of abusive language in students' open-ended comments is aimed at female and other minority instructors. Survey evidence indicates that about 52% of males report having received abusive comments, compared to 63% of women and 60% of non-binary individuals [24]. Additionally, while 55% (60%) of straight men (straight women) report having received abusive comments, 64% (83%) of gay (lesbian) instructors report having been on the receiving end of such abuse [24]. Lastly, as Heffernan [24] also points out, exploration by DiPietro and Faye [28] and Hamermesh and Parker [29] finds at least limited evidence of traditional SET discrimination against instructors who are visibly disabled, or who are viewed by students as not being heterosexual or a binary gender.

### 2.3. Grading Effect in SETs

Krautmann and Sander [30] expertly qualify the problems with SETs in relation to students' grades in economics. As they point out, if SETs can be improved through the assignment of higher grades, then they are a flawed instrument for the evaluation of teaching. This flaw may be contributing to grade inflation, unmeritorious decisions regarding tenure and promotion, and a dilution of the signaling role of educational credentials in screening workers for the labor market [30]. Regression results presented in both Krautmann and Sander [30] and the aforementioned studies by McPherson [10] and McPherson et al. [11] support the latter's explanations in finding a positive and significant relationship between SETs and students' grade expectations. In fact, all three studies conclude that instructors can "buy" better SET scores by inflating students' grade expectations. Lastly, a more recent study by Matos-Díaz and Ragan [31] asserts that, because of risk aversion, SETs are dependent upon the characteristics of the distribution of class grades. Matos-Díaz and Ragan [31] find support for this assertion in their own analysis of SETs from the University of Puerto Rico, which indicates that SETs are significantly and negatively related to the variance of expected grades, implying that faculty may be able to boost their SETs by narrowing the grade distribution, particularly in the case of the weakest students.

### 2.4. Experience Effect in SETs

Although prior research reports that SET scores are not impacted by an instructor's experience or academic rank [12], these potential determinants have remained a focus of subsequent studies. For example, McPherson [10] employs a longitudinal approach in examining 607 economics classes over 17 semesters in order to account for unobserved heterogeneity. In the case of economics principles classes, McPherson [10] and McPherson et al. [11] find that the level of experience of the instructor is a significant determinant of SET ratings. A more recent study by Alauddin and Kifle [32] applies SET data from an elite Australian university to a partial proportional odds model to investigate the influence of students' perceptions of instructional attributes on SETs. Among its many findings, the study reports that instructors below the rank of associate professor earned higher SETs [32]. From a U.S. perspective, the highest performers would include instructors, lecturers and assistant professors. Lastly, Keng's [16] examination of SETs from a public university in Taiwan that relies upon a learning model to measure teaching effectiveness finds that instructor experience is positively related to teaching effectiveness. More specifically, the study reports that the gender bias in teaching evaluations is reduced by nearly 50% after 10 years of teaching [16].

### 2.5. Efficacy of RMP Data in Academic Research

Low response rates associated with SETs, particularly those conducted online, often raise questions about the validity of RMP data as they could suffer from non-response bias [33]. To address this issue, Layne et al. [34] examine SET data across five academic disciplines from two groups of students ( $n = 2453$ ) at a large university, one of which completed paper-and-pencil SET surveys while the other utilized an electronic mode. Statistical analysis discussed in the study indicated that response rates differ by mode of administration. Even so, the actual SET ratings were not significantly influenced by the survey method, suggesting that the electronic survey mode is a viable alternative to the paper-and-pencil mode of administration [34]. A similar examination by Avery et al. [35] of SET data from a large economics-based public policy program at Cornell University revealed that although Web-based evaluation methods lead to lower response rates, these lower response rates do not appear to impact mean evaluation scores. This result suggests that SET ratings are not adversely affected by switching from paper to online evaluations [35]. Lastly, a more recent study by Nowell et al. [36] uses the Heckman [37] two-step selection correction procedure to account for potential sample selection bias in online SETs and finds no existence of such bias. Tangential research by Bleske-Rechek and Michels [38] examines data on the use of RMP from 208 students at a regional public

university and finds that the characteristics of students who tend to post ratings on RMP do not differ from those who do not post on the website.

A number of academic studies compare the ratings of faculty across in-house SETs and those posted to RMP. For example, studies by Coladarci and Kornfield [39], Timmerman [40], Albrecht and Hoopes [41], Brown et al. [42] and Sonntag et al. [43] compare RMP ratings with official institution-administered SETs and conclude that ratings from the two delivery modes are highly correlated. More specifically, Coladarci and Kornfield [39] compare SET data for both an in-house instrument and RMP across 426 instructors at the University of Maine and find a correlation coefficient of 0.68. Timmerman [40] compares similar data on instructors from the University of Tennessee and the University of Colorado, Boulder, and finds correlation coefficients of 0.67 and 0.77, respectively. Albrecht and Hoopes [41] compare RMP data to official evaluation data on 243 faculty from the business schools of one large private research-oriented university and one large public teaching-oriented university in the U.S. They find correlations ranging from 0.62 to 0.81, suggesting that students could legitimately use RMP data to compare different professors within a university. A reliability analysis of RMP and SET ratings of 312 faculty at Brooklyn College by Brown et al. [42] revealed strong internal consistency between the two evaluation formats. In terms of overall quality, the two formats produced a correlation coefficient of 0.64 [42]. Sonntag et al.'s [43] comparison of RMP and in-house SET ratings of 126 faculty at Lander University produces a correlation coefficient of 0.69. This, they argue, indicates that the quality ratings from RMP provide students with information about instructors that is comparable to the information they would have if institutionally administered evaluations were made available to them [43]. Lastly, related research by Otto et al. [44] analyzes the pattern of relationships of RMP ratings for 399 randomly selected faculty. Their results indicate that the pattern of RMP ratings is consistent with the pattern expected of a valid measure of student learning [44].

### *2.6. Use of RMP Data in Academic Research*

RMP provides the largest publicly available online SET data source, and it has been widely employed in the business education literature (e.g., [45,46]). Many studies discuss straightforward analyses of the determinants of RMP ratings. Constand and Pace [47], for example, compare RMP evaluation scores of finance faculty to those of other disciplines, finding that perceived course difficulty is an important determinant of students' ratings. Boehmer and Wood [48] examine RMP data to explore the role of faculty gender and course rigor on students' ratings of faculty. Their results indicate that students prefer male instructors and instructors who offer less rigorous courses, as these instructors receive the highest RMP scores [48]. Similarly, Constand et al. [49] report that accounting students perceive their professors to be significantly more difficult than students in non-accounting disciplines and this perception is related to lower teaching evaluations. More recently, Constand and Clarke [50] and Constand et al. [51] examined a large sample of RMP instructor ratings for undergraduate courses taught at nine Florida universities and found that instructors who teach introductory/core courses earn lower ratings than instructors who teach either advanced or very advanced/capstone courses, and that instructors who teach advanced courses earn lower ratings than those who teach very advanced/capstone courses. In other words, these studies find a cascading effect where instructor ratings increase as course level increases.

Carter [52] investigates the notion that research-active faculty offer superior instruction. Controlling for faculty gender and rank, tuition, and other student and institution effects, results suggest that faculty scholarship is positively related to RMP ratings only for male faculty who publish in elite journals [52]. A number of prior studies have used information on instructor attractiveness found in previous iterations of the RMP platform. Smith [53], for example, finds that relatively more attractive instructors, a notion captured by the proportion of an instructor's RMP ratings that include a designation indicating that the instructor is viewed by the raters as being physically attractive, earn RMP ratings that are

significantly greater than their counterparts, holding constant course rigor, institutional type and selectivity, and tuition. Mixon and Smith [54] extend this analysis by showing that relatively more attractive instructors tend to offer more rigorous courses, according to RMP data, than their counterparts. They argue that this result suggests that by trading on their attractiveness in offering more rigorous courses as opposed to using their attractiveness as a supplement to offering a less rigorous course, relatively attractive faculty are able to maintain a relatively good standing amongst their departmental or unit peers [54]. Finally, studies by Green et al. [55,56] use RMP data on instructor attractiveness to explain self-sorting in higher education. More specifically, these studies find that relative attractiveness is a significant predictor of institutional choice in higher education, whereby attractive faculty are more likely to choose employment at liberal arts institutions, where teaching is the primary responsibility, in order to capitalize on their looks [55,56].

### 3. Econometric Model

In order to better understand how the quality of instruction provided by clinical faculty in economics compares to that provided by traditional tenure-track economics faculty, we propose the double-log econometric model specification,

$$\ln TeachQual_i = \alpha + \beta_1 Female_i + \beta_2 FullProf_i + \beta_3 AssocProf_i + \beta_4 ClinicalProf_i + \beta_5 Private_i + \beta_6 \ln Facilities_i + \beta_7 \ln Internet_i + \beta_8 \ln MedSAT_i + \beta_9 \ln SFRatio_i + \beta_{10} \ln DiffCourse_i + \varepsilon, \quad (1)$$

where the dependent variable,  $\ln TeachQual_i$ , is the log of the mean of faculty  $i$ 's teaching quality ratings. The RMP scale for teaching quality is: 5 = awesome, 4 = great, 3 = good, 2 = OK, and 1 = awful. Individual student ratings based on this scale are averaged for each instructor, and this average is provided by RMP. As indicated in (1) above,  $\ln TeachQual_i$  is partially determined by  $Female_i$ , which is a dummy variable equal to 1 if faculty  $i$  is female, and 0 otherwise. Recent research [13–17] consistently indicates that female faculty garner lower ratings of teaching quality than their male counterparts. Thus, we expect that the regression estimate for  $\beta_1$  (i.e.,  $b_1$ ) will be negatively signed.

Next, a dummy variable series for academic rank is included on the right-hand side of (1). Included in this series are  $FullProf_i$  and  $AssocProf_i$ . The first of these is a dummy variable equal to 1 if faculty  $i$  holds the rank of full professor and 0 otherwise. The second of these is a dummy variable equal to 1 if faculty  $i$  holds the rank of associate professor and 0 otherwise. According to research by Liaw and Goh [12], SET scores are not impacted by an instructor's experience or academic rank. If so, estimates of  $\beta_2$  and  $\beta_3$  (i.e.,  $b_2$  and  $b_3$ ) will not significantly differ from zero. On the other hand, if the finding in McPherson [10], McPherson et al. [11] and Keng [16] that teaching experience matters is correct, then  $b_2$  and  $b_3$  will be positively signed. Negatively signed estimates in this case would, on the other hand, support research by Alauddin and Kifle [32].

Some institutional variables are also included on the right-hand side in (1). The first of these,  $Private_i$ , is a dummy variable equal to 1 for private colleges and universities, and 0 otherwise. The colleges and universities that reside in this category exist along a continuum where at one end are institutions that maintain a liberal arts focus and attract students who desire a well-rounded education that is provided in a teaching-centered environment, while at the other end are elite research institutions that expect faculty to publish in top academic journals. As such, we make no a priori assertion regarding the sign of the parameter estimate attached to  $\beta_5$  (i.e.,  $b_5$ ). Next, research by Osoian et al. [57] and Benton and Cashin [58] finds that factors such as a university's classroom design, cleanliness, website quality, library services, and food options may influence instructor SETs even more so than its teaching standards. As such, our model includes similar variables. For example,  $\ln Facilities_i$  is the log of the mean of students' RMP ratings of the academic facilities comprising the institution employing faculty  $i$ , while  $\ln Internet_i$  is the log of the mean of students' RMP ratings of the internet infrastructure available at the institution employing faculty  $i$ . If high-quality facilities and internet service enhance students' educational experiences, as one might expect, econometric estimates of  $\beta_6$  and  $\beta_7$

(i.e.,  $b_6$  and  $b_7$ ) will likely be positively signed. Next,  $\ln MedSAT_i$ , is the log of the median SAT score for incoming freshmen at the institution employing faculty  $i$ . To the extent that high-achieving students are more discerning or discriminating with regard to relatively lower quality instruction, the estimate of  $\beta_8$  (i.e.,  $b_8$ ) will be negatively signed. The final institutional variable,  $\ln SFRatio_i$ , is the log of the student-to-faculty ratio at the institution employing faculty  $i$ . The impersonal nature of the educational experience in academic settings where the student-to-faculty ratio is relatively high is likely to be associated with lower ratings of teaching quality, ceteris paribus. As such, we expect the econometric estimate of  $\beta_9$  (i.e.,  $b_9$ ) to be negatively signed.

Research by Krautmann and Sander [30], McPherson [10], McPherson et al. [11] and Matos-Díaz and Ragan [31] that is reviewed in the prior section suggests that students' grade expectations are directly related to their ratings of instructors' performance. Given that students tend to negatively associate course difficulty with expected grades, one would expect course difficulty to be inversely related to SETs. As such, the regression specification in (1) above includes  $\ln DiffCourse_i$ , which represents the log of the mean difficulty rating for each faculty,  $i$ . The RMP scale for course difficulty is: 5 = very difficult, 4 = difficult, 3 = average, 2 = easy, and 1 = very easy. Individual student ratings based on this scale are averaged for each instructor, and this average is provided by RMP. The regression estimate of  $\beta_{10}$  (i.e.,  $b_{10}$ ) is expected to be negatively signed. Lastly, the variable of interest in (1),  $ClinicalProf_i$ , is a dummy variable equal to 1 if faculty  $i$  holds a clinical position and 0 otherwise. Given the specialized teaching nature of clinical faculty positions, one would expect that econometric estimation of (1) would produce a positively-signed estimate of  $\beta_4$  (i.e.,  $b_4$ ), ceteris paribus. This possibility is explored further in the latter part of this section of the study.

#### 4. Data

In collecting data for this study, we examined the composition of economics departments at all institutions in the national colleges and universities category of *U.S. News & World Report's* annual guide, *America's Best Colleges*. In doing so, we discovered the presence of clinical faculties in the economics departments at each of the national colleges and universities listed in Table A1 of Appendix A. In identifying the clinical faculty at the institutions in Table A1 using faculty profiles on department webpages, we sought a combination of terms including "professor" and either "clinical", "practice", "teaching", or "instruction", or at least some form of one of the latter four terms. The only exception we entertained involved what appeared to be named professorships in the clinical realm. Of the eight of these that we found, two did not include the term "professor" in their names. As stated above, teaching quality (SET) data are collected from RateMyProfessors.com (RMP) for the economics faculty affiliated with the universities listed in Table A1 who have five or more RMP ratings. Data for both  $Facilities_i$  and  $Internet_i$  are also collected from RMP. Lastly, data for  $MedSAT_i$  and  $SFRatio_i$  are collected from *America's Best Colleges*.

Variable names, descriptions and summary statistics are presented in Table 1. As indicated there, the mean of all 947 teaching quality ratings means is 3.419 (out of 5). Next, 26.5% of the sample is represented by female faculty, while clinical faculty constitute 24.6% of the sample. The breakdown by professorial rank indicates that 51.1% of the sample is represented by full professors, 30.1% by associate professors and 18.8% by assistant professors. Just under 14% of professors in the sample hold named professorships, while 28% are employed by private universities. In terms of the institutional and student variables, the average of all university facilities' quality ratings means is 4.154 (out of 5), while the average of all university internet quality ratings means is 3.713 (out of 5). The average student-to-faculty ratio is 16.2 students, the mean percentage of classes with fewer than 20 students across the entire sample is 44.5%, and the average of all course difficulty ratings means is 3.374 (out of 5). Lastly, although not shown in Table 1, when the sample is divided across professorial category, the mean of  $TeachQual_i$  for clinical faculty is 3.600, while that

for all others (i.e., traditional tenure-track faculty) is 3.360. The difference, or 0.240, is based on a standard error of 0.062, greater than 0 at better than the 0.01 level of significance.

**Table 1.** Variable descriptions and summary statistics.

| Variable                        | Variable Description   | Mean   | Std. Dev. | N   |
|---------------------------------|--|--------|-----------|-----|
| <i>TeachQual<sub>i</sub></i>    | Mean of teaching quality ratings (1 to 5) for each faculty, <i>i</i> .                                     | 3.419  | 0.831     | 947 |
| <i>Female<sub>i</sub></i>       | Dummy variable equal to 1 if faculty <i>i</i> is female, and 0 otherwise.                                  | 0.265  | 0.442     | 947 |
| <i>FullProf<sub>i</sub></i>     | Dummy variable equal to 1 if faculty <i>i</i> is a full professor, and 0 otherwise.                        | 0.511  | 0.500     | 947 |
| <i>AssocProf<sub>i</sub></i>    | Dummy variable equal to 1 if faculty <i>i</i> is an associate professor, and 0 otherwise.                  | 0.301  | 0.459     | 947 |
| <i>ClinicalProf<sub>i</sub></i> | Dummy variable equal to 1 if faculty <i>i</i> is classified as a clinical professor, and 0 otherwise.      | 0.246  | 0.431     | 947 |
| <i>NamedProf<sub>i</sub></i>    | Dummy variable equal to 1 if faculty <i>i</i> holds a named professorship, and 0 otherwise.                | 0.137  | 0.344     | 947 |
| <i>Private<sub>i</sub></i>      | Dummy variable equal to 1 if faculty <i>i</i> is employed by a private university, and 0 otherwise.        | 0.280  | 0.449     | 947 |
| <i>Facilities<sub>i</sub></i>   | Mean of facilities quality ratings (1 to 5) of university employing each faculty, <i>i</i> .               | 4.154  | 0.354     | 947 |
| <i>Internet<sub>i</sub></i>     | Mean of internet quality ratings (1 to 5) of university employing each faculty, <i>i</i> .                 | 3.713  | 0.306     | 947 |
| <i>MedSAT<sub>i</sub></i>       | Median SAT score for incoming freshman of university employing each faculty, <i>i</i> .                    | 1299.8 | 120.7     | 936 |
| <i>SFRatio<sub>i</sub></i>      | Student-to-faculty ratio of university employing each faculty, <i>i</i> .                                  | 16.23  | 4.761     | 939 |
| <i>PctFew20<sub>i</sub></i>     | Percentage of classes with fewer than 20 students offered by university employing each faculty, <i>i</i> . | 44.50  | 13.18     | 939 |
| <i>DiffCourse<sub>i</sub></i>   | Mean of course difficulty ratings (1 to 5) for each faculty, <i>i</i> .                                    | 3.374  | 0.596     | 947 |

Table 2 provides a correlation matrix including both *TeachQual<sub>i</sub>* and its determinants. Although many of the correlation coefficients in Table 2 are significantly different from 0, most are relatively small (i.e., near 0). Some of the significant correlations are, however, quite interesting. First, there is a significant positive correlation between *Female<sub>i</sub>* and *ClinicalProf<sub>i</sub>*, suggesting an association between gender and faculty status. Second, there is a significant negative association between *ClinicalProf<sub>i</sub>* and *DiffCourse<sub>i</sub>*, suggesting that clinical faculty tend offer less difficult economics courses than their traditional tenure-track faculty counterparts. This raises the question of whether the courses offered by clinical faculty are structured to be less taxing on students, or that clinical faculty possess instructional skills that make their courses appear easier to economics students. Relatedly, Table 2 also reveals significantly negative correlations between *Female<sub>i</sub>* and *FullProf<sub>i</sub>*, and between *Female<sub>i</sub>* and *NamedProf<sub>i</sub>*. These suggest that males have a significant advantage in both rank and named professorship attainment. These significantly negative correlation coefficients involving *Female<sub>i</sub>* and both *NamedProf<sub>i</sub>* and *FullProf<sub>i</sub>* are consistent with studies by Sabatier [59], Cooray et al. [60], Bukstein and Gandelman [61] and Thorndyke et al. [62].

**Table 2.** Correlation matrix.

|                     | <i>TeachQual</i> | <i>Female</i> | <i>FullProf</i> | <i>AssocProf</i> | <i>ClinicalProf</i> | <i>NamedProf</i> | <i>Private</i> | <i>Facilities</i> | <i>Internet</i> | <i>MedSAT</i> | <i>SFRatio</i> | <i>PctFew20</i> |
|---------------------|------------------|---------------|-----------------|------------------|---------------------|------------------|----------------|-------------------|-----------------|---------------|----------------|-----------------|
| <i>Female</i>       | −0.011           |               |                 |                  |                     |                  |                |                   |                 |               |                |                 |
| <i>FullProf</i>     | −0.033           | −0.222        |                 |                  |                     |                  |                |                   |                 |               |                |                 |
| <i>AssocProf</i>    | +0.032           | +0.107        | −0.671          |                  |                     |                  |                |                   |                 |               |                |                 |
| <i>ClinicalProf</i> | +0.125           | +0.212        | −0.197          | +0.037           |                     |                  |                |                   |                 |               |                |                 |
| <i>NamedProf</i>    | −0.016           | −0.128        | +0.341          | −0.208           | −0.178              |                  |                |                   |                 |               |                |                 |
| <i>Private</i>      | +0.047           | −0.039        | +0.087          | −0.014           | +0.004              | +0.032           |                |                   |                 |               |                |                 |
| <i>Facilities</i>   | −0.057           | −0.033        | +0.082          | −0.076           | −0.030              | +0.063           | +0.143         |                   |                 |               |                |                 |
| <i>Internet</i>     | −0.048           | −0.057        | +0.063          | −0.039           | +0.025              | +0.067           | +0.189         | +0.560            |                 |               |                |                 |
| <i>MedSAT</i>       | −0.007           | −0.015        | +0.174          | −0.109           | −0.013              | +0.112           | +0.566         | +0.270            | +0.415          |               |                |                 |
| <i>SFRatio</i>      | −0.029           | +0.041        | −0.118          | +0.051           | −0.008              | −0.027           | −0.765         | −0.222            | −0.153          | −0.645        |                |                 |
| <i>PctFew20</i>     | +0.021           | −0.056        | +0.123          | −0.057           | −0.020              | +0.129           | +0.653         | +0.207            | +0.207          | +0.612        | −0.674         |                 |
| <i>DiffCourse</i>   | −0.446           | −0.040        | −0.011          | −0.009           | −0.131              | +0.002           | −0.006         | +0.083            | −0.002          | +0.018        | −0.030         | −0.017          |

## 5. Econometric Results and Discussion

### 5.1. Individual Level Results

Results from ordinary least squares (OLS) estimation of (1) are presented in Table 3. As indicated in the second column of the table, the regressors included in (1) are jointly significant (at the 0.01 level) in explaining the variation in *lnTeachQual*, while they account for 19% of that variation. The coefficient estimate attached to *Female* suggests that female faculty earn teaching quality ratings that are 2.8% lower than those of their male counterparts, ceteris paribus. However, this result is marginally insignificant at the usual levels.

Interestingly, the negatively signed coefficient estimate attached to  $\ln Internet$  is statistically significant, suggesting that where internet service is perceived by students to be good, teaching quality ratings of economics faculty are lower. More specifically, a 10% increase in perceived internet service quality is associated with teaching quality ratings that are about 1.9% lower, *ceteris paribus*. The first set of regression results in Table 3 also suggests that course difficulty is penalized by student raters. In this case, a 10% increase in course difficulty is associated with a 5.7% decrease in the teaching quality ratings of economics instructors, *ceteris paribus*. Lastly, the parameter estimate attached to our variable of interest, *ClinicalProf*, is both positively signed and statistically significant. It suggests that clinical economics faculty earn teaching quality ratings that are 4.4% higher than those of their traditional tenure-track economics faculty counterparts, *ceteris paribus*.

**Table 3.** Individual-level results.

| Regressors                             | (1)                  | (2)                  | (3)                  | (4)                  |
|--|----------------------|----------------------|----------------------|----------------------|
| constant                               | 2.875 *<br>(3.33)    | 2.532 *<br>(3.40)    | 1.802 *<br>(18.66)   | 1.776 *<br>(19.87)   |
| <i>Female</i>                          | −0.028<br>(−1.45)    | −0.029<br>(−1.48)    | −0.035 †<br>(−1.90)  | −0.039 †<br>(−2.06)  |
| <i>FullProf</i>                        | 0.004<br>(0.15)      | 0.004<br>(0.15)      | −0.009<br>(−0.36)    | −0.011<br>(−0.46)    |
| <i>AssocProf</i>                       | 0.020<br>(0.76)      | 0.020<br>(0.76)      | 0.008<br>(0.31)      | 0.006<br>(0.25)      |
| <i>ClinicalProf</i>                    | 0.043 †<br>(2.11)    | 0.043 †<br>(2.14)    | 0.047 †<br>(2.25)    | 0.038 †<br>(1.81)    |
| <i>NamedProf</i>                       | −0.004<br>(−0.16)    | −0.004<br>(−0.14)    | −0.003<br>(−0.13)    | −0.015<br>(−0.53)    |
| <i>NamedProf</i> × <i>ClinicalProf</i> |                      |                      |                      | 0.201 *<br>(2.74)    |
| <i>Private</i>                         | 0.033<br>(1.17)      | 0.048 †<br>(2.16)    |                      |                      |
| $\ln Facilities$                       | 0.034<br>(0.33)      | 0.049<br>(0.48)      |                      |                      |
| $\ln Internet$                         | −0.204 †<br>(−1.67)  | −0.222 †<br>(−1.84)  |                      |                      |
| $\ln MedSAT$                           | −0.096<br>(−0.82)    | −0.047<br>(−0.41)    |                      |                      |
| $\ln SFRatio$                          | −0.022<br>(−0.51)    |                      |                      |                      |
| $\ln PctFew20$                         |                      | −0.018<br>(−0.51)    |                      |                      |
| $\ln CourseDiff$                       | −0.611 *<br>(−11.42) | −0.611 *<br>(−11.43) | −0.627 *<br>(−11.67) | −0.627 *<br>(−11.71) |
| Institution Effects                    | no                   | no                   | yes                  | yes                  |
| <i>n</i>                               | 936                  | 936                  | 947                  | 947                  |
| <i>F</i> -statistic                    | 19.65 *              | 19.65 *              | 3.78 *               | 3.79 *               |
| <i>R</i> <sup>2</sup>                  | 0.190                | 0.190                | 0.294                | 0.297                |

Notes: The numbers in parentheses above are robust *t*-ratios [63]. \*(†)[†] denotes the 0.01(0.05)[0.10] level of significance.

The second set of results in Table 3 involves a substitution of  $\ln PctFew20$  for  $\ln SFRatio$ , the latter of which is negatively associated with  $\ln TeachQual$  in the first set of results discussed above. Most of these results are similar to those from the previous regression, with the exception of *Private*, which is positively and significantly associated with  $\ln TeachQual$ , suggesting that private university students rate the teaching quality of their economics faculty relatively higher than their public university counterparts. In this case, economics

faculty who are affiliated with private universities earn teaching quality ratings that are about 4.9% higher than those of their public university counterparts, *ceteris paribus*. Lastly, the parameter estimate attached to our variable of interest, *ClinicalProf*, is again both positively signed and statistically significant, and of the same magnitude as that in the initial set of results.

### 5.2. Sensitivity Analysis

The robustness of the estimates above is examined by employing a propensity score analysis (PSA) approach, which relies on the estimation of the probability of receiving treatment, or propensity score [64,65]. For sensitivity analysis, in this study, we allow for counterfactual comparisons by using propensity score weighting (PSW), which employs all observations in the original sample but weights them according to their propensity scores [65]. As pointed out by Narita et al. [65], PSW uses the inverse of the propensity score as a weight to apply to each treated unit, and the inverse of one minus the propensity score as the weight to apply to each control unit [66]. PSA consists of estimating the impact of treatment on the variable of interest, which in our case is *ClinicalProf*. The treatment effect is obtained by comparing the average outcomes between treated and control units based on a comparison of the weighted average of outcomes between treated and control groups (i.e., PSW) using multiple regression [65]. The PSA approach used in this study provides the average treatment effect (ATE), which results from evaluating the impact of treatment on the whole weighted sample [65].

Table 4 provides estimates of the ATE coefficients for the clinical professor dummy based on the specifications in the second and third columns, respectively, of Table 3. These are interpreted in the same manner as the dummy variable coefficient estimates in Table 3. Thus, the first ATE coefficient estimate in Table 4 of 0.056 suggests that clinical faculty earn teaching quality ratings that are 5.8% higher than those of their tenure-track faculty counterparts, *ceteris paribus*. The second estimate in Table 4 of 0.059 similarly suggests that clinical faculty earn teaching quality ratings that are 6.1% higher than those of their tenure-track faculty counterparts, *ceteris paribus*. Each of these results is larger than its Table 3 counterpart and significant at about the 0.013 level, which is notably better than the results relating to *ClinicalProf* shown in Table 3.

**Table 4.** Average treatment effects, *ClinicalProf*.

| ATE                           | (1)               | (2)               |
|-------------------------------|-------------------|-------------------|
| <i>ClinicalProf</i> (1 vs. 0) | 0.056 ‡<br>(2.54) | 0.059 ‡<br>(2.54) |

Notes: The numbers in parentheses above are z-statistics based on robust standard errors. ‡ denotes the 0.05 level of significance.

### 5.3. Institutional Effects Estimates

The fourth column of Table 3 presents results from an alternative approach that avoids the troublesome multicollinearity between the institutional and student characteristics. This version employs a categorical set of dummy variables for the 90 institutions in the sample (see Table A1) instead of the institutional and student characteristics included in prior specifications. The omitted institution using this approach is the University of Alabama, Huntsville. As indicated in the fourth column of Table 3, this approach uses all 947 observations, is jointly significant and produces an  $R^2$  of 0.294. The coefficient estimate attached to *Female* suggests that female faculty earn teaching quality ratings that are 3.4% lower than those of their male counterparts, *ceteris paribus*. In this case, the result is statistically significant. This third set of regression results in Table 3 again suggests that course difficulty is penalized by student raters. In this case, a 10% increase in course difficulty is associated with a 5.8% decrease in the teaching quality ratings of economics faculty, *ceteris paribus*. Lastly, the parameter estimate attached to our variable of interest, *ClinicalProf*, is again positively signed and statistically significant. In this case, it suggests

that clinical faculty earn teaching quality ratings that are 4.8% higher than those of their tenure-track faculty counterparts, *ceteris paribus*.

The fifth column of Table 3 provides OLS estimation results of a specification identical to that from the fourth column, with the addition of an interaction term involving *NamedProf* and *ClinicalProf*. This specification is jointly significant and produces an  $R^2$  nearly identical to that from the prior specification. In this specification, the negatively-signed parameter estimate attached to female is significant at the 0.05 level, and indicates that female faculty earn teaching quality ratings that are 3.8% lower than those of their male counterparts, *ceteris paribus*. Next, the results regarding *DiffCourse* are in this case identical to those from the previous specification. That is, a 10% increase in course difficulty is associated with a 6.2% decrease in the teaching quality ratings of economics faculty, *ceteris paribus*. Lastly, the results concerning *NamedProf* and *ClinicalProf* are quite interesting. They indicate that (1) unnamed clinical faculty earn teaching quality ratings that are 3.9% higher than those of their unnamed traditional tenure-track counterparts, (2) named traditional tenure-track faculty earn teaching quality ratings that are 1.5% lower than those of their unnamed traditional tenure-track counterparts, and (3) named clinical faculty earn teaching quality ratings that are 25.1% greater than those of their unnamed traditional tenure-track counterparts, *ceteris paribus*.

#### 5.4. Department Level Results

Our final exercise explores the magnitude of the teaching ethos in those economics departments in our sample that employ clinical faculty. This is carried out by the creation of the variable  $DeptTQual_j$ , which is equal to the average of the mean teaching quality ratings across the faculty in economics department  $j$ . This variable is first used to rank the top 30 economics departments included in our sample. This ranking is shown in Table 5. As indicated in Table 5, the economics department at the University of California, Merced, sits atop the ranking with a  $DeptTQual$  of 4.267 (out of 5), followed by Boston College, the University of Florida, the University of Illinois, Chicago, and the University of Texas, Arlington. Entry into the top five in Table 5 requires a departmental average teaching quality (i.e.,  $DeptTQual_j$ ) of 3.920 or better. The second half of the top 10 includes the University of Massachusetts, Lowell, Rutgers University, Camden, Seattle University, the University of Wisconsin, La Crosse, the University of Miami and the University of San Diego. Entry into this group requires a departmental average teaching quality of 3.810 or better.

**Table 5.** Top 30 economics departments by departmental teaching quality.

| Rank | Institution                         | DeptTQual | Rank | Institution                             | DeptTQual |
|------|-------------------------------------|-----------|------|---|-----------|
| 1    | University of California, Merced    | 4.267     | 16   | Florida State University                | 3.671     |
| 2    | Boston College                      | 4.050     | 17   | Northwestern University                 | 3.650     |
| 3    | University of Florida               | 4.025     | 18   | University of Rhode Island              | 3.644     |
| 4    | University of Illinois, Chicago     | 3.944     | 19   | Tulane University                       | 3.640     |
| 5    | University of Texas, Arlington      | 3.920     | 20   | University of Houston                   | 3.621     |
| 6    | University of Massachusetts, Lowell | 3.914     | 21   | University of Nebraska                  | 3.613     |
| 7    | Rutgers University, Camden          | 3.900     | 22   | University of Memphis                   | 3.600     |
|      | Seattle University                  | 3.900     |      | University of North Carolina, Charlotte | 3.600     |
| 9    | University of Wisconsin, La Crosse  | 3.850     | 24   | DePaul University                       | 3.594     |
| 10   | University of Miami                 | 3.810     | 25   | Villanova University                    | 3.591     |
|      | University of San Diego             | 3.810     | 26   | University of Texas, San Antonio        | 3.588     |
| 12   | Purdue University                   | 3.792     | 27   | Ball State University                   | 3.582     |
| 13   | Lehigh University                   | 3.713     | 28   | University of Chicago                   | 3.567     |
| 14   | University of Texas                 | 3.693     | 29   | Rutgers University                      | 3.550     |
| 15   | Washington University, St. Louis    | 3.673     | 30   | Wake Forest University                  | 3.533     |

Further examination of Table 5 reveals that 40% of economics departments that exhibit a relatively strong teaching ethos are affiliated with private universities. This group includes institutions where research plays a major role (e.g., Northwestern University), as well as institutions where teaching is the primary focus (e.g., University of San Diego). A similar

mix of public universities is also included among the top 30 institutions listed in Table 5. Rutgers University, Camden, for example, represents a public university where teaching is prominent, while the University of Texas provides a good example of a public university where high-quality research is expected.

In order to better understand the relationship between the quality of instruction and the deployment of clinical faculty at the departmental level, we regressed  $\ln DeptTQual_j$ , or the logarithm of  $DeptTQual_j$ , for the top 30 economics departments listed in Table 5 on  $\ln FacultyRank_j$ ,  $\ln ClinicalProfRatio_j$ ,  $Private_j$  and  $\ln DCourseDiff_j$ . The first of these variables represents the logarithm of the mean academic rank of the faculty in economics department  $j$ , whereby assistant professors are coded as 1, associate professors are coded as 2, and professors are coded as 3. The variable of interest,  $\ln ClinicalProfRatio_j$ , is the logarithm of the ratio of clinical faculty to total faculty in economics department  $j$ . If clinical economics faculty provide superior economics instruction compared to traditional tenure-track faculty, as the results in Tables 3 and 4 suggest, then the deployment of a greater proportion of clinical faculty will improve average instructional quality in economics department  $j$  (i.e., parameter estimate attached to  $\ln ClinicalProfRatio_j$  should be positively signed). Lastly,  $Private_j$  is a dummy variable equal to 1 for economics departments that are attached to private universities, while  $\ln DCourseDiff_j$  is the logarithm of the mean course difficulty offered by the faculty of economics department  $j$ . As before, evidence that higher SETs can be purchased would be supported by a negatively signed and significant parameter estimate attached to  $\ln DCourseDiff_j$ .

OLS estimates for the department-level approach are presented in Table 6. As indicated in the second column of the table, the regressors included in this unrestricted specification are jointly significant (at the 0.01 level) in explaining the variation in  $\ln DeptTQual$ , while they account for about 51% of that variation. To begin, the positively signed coefficient estimate attached to  $Private$  is near zero and not statistically significant. Next, the negatively signed coefficient estimate attached to  $\ln FacultyRank$  suggests that economics departments with faculty who, on average, hold higher academic ranks produce, on average, lower teaching quality ratings. More specifically, a 10% increase in the average academic rank of an economics department's faculty is associated with departmental teaching quality ratings that are about 0.8% lower, ceteris paribus. The first set of regression results in Table 3 also suggests that course difficulty is penalized by student raters. In this case, a 10% increase in mean departmental course difficulty is associated with a 3.1% decrease in the teaching quality ratings of economics instructors, ceteris paribus. Lastly, the parameter estimate attached to our variable of interest,  $\ln ClinicalProfRatio$ , is both positively signed and statistically significant. In this case, a 10% increase in the ratio of clinical faculty to total faculty in economics department  $j$  leads to only a 0.3% increase in average departmental teaching quality, ceteris paribus. In the context of this model, however, a 10% increase in the ratio of clinical faculty is achieved when economics department  $j$ 's deployment of clinical faculty rises from 1 (out of 10) to 1.1, a change that is meaningless. Of course, it is possible for economics department  $j$  to increase the deployment of clinical faculty from 1 (out of 10) to 2. In this case, one would expect to observe a 2.2% increase in average departmental teaching quality, ceteris paribus.

The additional results in Table 6 come from restricted versions of the first model. Here, the parameter estimate attached to  $\ln DCourseDiff$  is  $-0.364$ . In this case, a 10% increase in mean departmental course difficulty is associated with a 3.4% decrease in the average teaching quality ratings of its economics instructors, ceteris paribus. Lastly, the parameter estimate attached to  $\ln ClinicalProfRatio$  climbs to 0.035. Thus, if economics department  $j$ 's deployment of clinical faculty rises from 1 (out of 10) to 2, one would expect to observe a 2.5% increase in average departmental teaching quality, ceteris paribus.

**Table 6.** Department-level results.

| Regressors                 | (1)                 | (2)                 | (3)                 | (4)               |
|----------------------------|---------------------|---------------------|---------------------|-------------------|
| constant                   | 1.812 *<br>(21.4)   | 1.798 *<br>(21.9)   | 1.801 *<br>(19.0)   | 1.798 *<br>(19.4) |
| <i>lnFacultyRank</i>       | −0.079<br>(−1.31)   |                     | −0.068<br>(−1.34)   |                   |
| <i>lnClinicalProfRatio</i> | 0.032 *<br>(2.91)   | 0.035 *<br>(2.94)   | 0.033 *<br>(2.98)   | 0.035 *<br>(2.94) |
| <i>Private</i>             | 0.007<br>(0.53)     | 0.000<br>(0.00)     |                     |                   |
| <i>lnDCourseDiff</i>       | −0.327 *<br>(−4.36) | −0.364 *<br>(−6.06) | −0.321 *<br>(−4.00) |                   |
| <i>n</i>                   | 30                  | 30                  | 30                  | 30                |
| <i>F</i> -statistic        | 6.62 *              | 8.11 *              | 9.02 *              | 12.63 *           |
| <i>R</i> <sup>2</sup>      | 0.514               | 0.483               | 0.510               | 0.483             |

Notes: The numbers in parentheses above are robust *t*-ratios [63]. \* denotes the 0.01 level of significance.

## 6. Conclusions

This analysis consistently documents higher student teaching quality ratings clinical economics faculty than for traditional tenure-track economics faculty. This difference, which is as large as 6.1%, is robust to different econometric model specifications. Moreover, we document a very large, 25.1%, student teaching quality rating premium for named clinical economics faculty relative to unnamed traditional tenure-track economics faculty. These findings strongly suggest that a deliberate search for high-quality teaching faculty pays off in terms of higher student evaluations of teaching. Lastly, our analysis yields an institutional ranking of economics departments based on the quality of their teaching. The presence of both public and private colleges and universities among the top 30 economics departments shows that the management strategy of hiring clinical faculty along with traditional tenure-track faculty can be successful in a variety of institutional settings. Overall, our findings are encouraging signs for the hiring and retention of clinical faculty in economics departments.

The multi-institutional approach to comparing the teaching quality ratings of clinical and traditional tenure-track faculty carries with it some inherent limitations. For example, we cannot account for the gender of the students issuing the ratings or for their grades. However, the large sample size and heterogeneous characteristics of the institutions and instructors included in our sample likely minimize any potential source of bias from those unobserved variables. Finally, there are multiple facets relating to how clinical faculty differ from traditional tenure-track faculty that are not explored in this study. These might include differences in educational training, compensation, course loads taught, types of teaching experience and professional background, to name a few. In fact, a formal exploration of these facets as a standalone study would be a useful contribution to the literature. Relatedly, given that the national institutions listed by *U.S. News & World Report* represent a unique population, future research might also examine how teaching quality provided by clinical and traditional tenure-track faculty varies across types of institutions, including across the multiple institutional categories discussed by *U.S. News & World Report*.

**Author Contributions:** Conceptualization, F.M.J.; methodology, S.C. and F.M.J.; formal analysis, S.C. and F.M.J.; data curation, F.M.J.; writing—original draft preparation, J.B., F.C. and F.M.J.; writing—review and editing, J.B., F.C. and F.M.J.; project administration, J.B. and F.M.J. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** Data are available from the authors upon request.

**Conflicts of Interest:** The authors have no conflicts of interests.

## Appendix A

**Table A1.** National universities utilizing clinical economics faculty.

| University of Alabama, Huntsville   | Georgia Institute of Technology         | University of Pittsburgh           |
|-------------------------------------|---|------------------------------------|
| Arizona State University            | Harvard University                      | Purdue University                  |
| University of Arkansas              | Johns Hopkins University                | Quinnipiac University              |
| Ball State University               | University of Houston                   | University of Rhode Island         |
| Baylor University                   | University of Illinois                  | University of Rochester            |
| Boston College                      | University of Illinois, Chicago         | Rutgers University                 |
| Bowling Green State University      | Iowa State University                   | Oklahoma State University          |
| University at Buffalo               | University of Kansas                    | Rutgers University, Camden         |
| University of California, Berkeley  | Kansas State University                 | University of San Diego            |
| University of California, Davis     | University of Kentucky                  | Seattle University                 |
| University of California, Irvine    | Kennesaw State University               | University of South Carolina       |
| University of California, Merced    | Lehigh University                       | University of South Florida        |
| University of California, Riverside | University of Massachusetts, Lowell     | Syracuse University                |
| University of California, San Diego | University of Memphis                   | Temple University                  |
| Carnegie Mellon University          | University of Miami                     | University of Texas                |
| University of Chicago               | University of Mississippi               | University of Texas, Arlington     |
| Clemson University                  | Mississippi State University            | University of Texas, El Paso       |
| University of Colorado              | University of Missouri                  | University of Texas, San Antonio   |
| University of Colorado, Denver      | University of Nebraska                  | Texas A&M University               |
| Colorado School of Mines            | University of Nevada, Reno              | Tufts University                   |
| University of Denver                | New York University                     | Tulane University                  |
| DePaul University                   | University of North Carolina            | Villanova University               |
| Drexel University                   | University of North Carolina, Charlotte | Wake Forest University             |
| Duke University                     | North Carolina State University         | University of Washington           |
| Emory University                    | Northeastern University                 | Washington University, St. Louis   |
| University of Florida               | Northwestern University                 | Wichita State University           |
| Florida International University    | University of Notre Dame                | University of Wisconsin, La Crosse |
| Florida State University            | University of Oregon                    | University of Wisconsin, Oshkosh   |
| Georgetown University               | Pennsylvania State University           | Worcester Polytechnic Institute    |
| Georgia State University            | Pepperdine University                   | Xavier University                  |

## References

- Pieters, G.; Roark, C. The job market for non-tenure track academic economists. In *FOCUS: A Guide for Non-Tenure Track Faculty*; Shreyasee, D., Seth, G., Eds.; CSWEP News, American Economic Association: Nashville, TN, USA, 2022.
- Mixon, F.G., Jr.; Upadhyaya, K.P. When forgiveness beats permission: Exploring the scholarly ethos of clinical faculty in economics. *Am. J. Econ. Sociol.* **2024**, *83*, 75–91.
- Asali, M. A tale of two tracks. *Educ. Econ.* **2019**, *27*, 323–337.
- Allen, M.; Sweeney, C. Faculty research productivity under alternative appointment types: Tenure vs non-tenure track. *Manag. Financ.* **2017**, *43*, 1348–1357. [\[CrossRef\]](#)
- Ran, F.X.; Xu, D. Does contractual form matter? The impact of different types of non-tenure-track faculty on college students' academic outcomes. *J. Hum. Resour.* **2019**, *54*, 1081–1120.
- Hilmer, C.; Hilmer, M. On the labor market for full-time non-tenure-track lecturers in economics. *Econ. Educ. Rev.* **2020**, *78*, 102023. [\[CrossRef\]](#)
- August, E.; Power, L.; Youatt, E.; Anderson, O. What does it mean to be a clinical track faculty member in public health? A survey of clinical track faculty across the United States. *Public Health Rep.* **2022**, *137*, 1235–1241. [\[CrossRef\]](#)
- Ongeri, J.D. Poor student evaluation of teaching in economics: A critical survey of the literature. *Australas. J. Econ. Educ.* **2009**, *6*, 1–24.
- Asarta, C.; Chambers, R.; Harter, C. Teaching methods in undergraduate introductory economics courses: Results from a sixth national quinquennial survey. *Am. Econ.* **2021**, *66*, 18–28.
- McPherson, M.A. Determinants of how students evaluate teachers. *J. Econ. Educ.* **2006**, *37*, 3–20. [\[CrossRef\]](#)
- McPherson, M.A.; Jewell, R.T.; Kim, M. What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *East. Econ. J.* **2009**, *35*, 37–51. [\[CrossRef\]](#)

12. Liaw, S.-H.; Goh, K.-L. Evidence and control of biases in student evaluations of teaching. *Int. J. Educ. Manag.* **2003**, *17*, 37–43.
13. Wagner, N.; Rieger, M.; Voorvelt, K. Gender, ethnicity and teaching evaluations: Evidence from mixed teaching teams. *Econ. Educ. Rev.* **2016**, *54*, 79–94. [[CrossRef](#)]
14. Boring, A. Gender biases in student evaluations of teaching. *J. Public Econ.* **2017**, *145*, 27–41. [[CrossRef](#)]
15. Mengel, F.; Sauermann, J.; Zölitz, U. Gender bias in teaching evaluations. *J. Eur. Econ. Assoc.* **2019**, *17*, 535–566. [[CrossRef](#)]
16. Keng, S.-H. Gender bias and statistical discrimination against female instructors in student evaluations of teaching. *Labour Econ.* **2020**, *66*, 101889.
17. Buser, W.; Hayter, J.; Marshall, E.C. Gender bias and temporal effects in standard evaluations of teaching. *AER Pap. Proc.* **2019**, *109*, 261–265. [[CrossRef](#)]
18. Chisadza, C.; Nicholls, N.; Yitbarek, E. Race and gender biases in student evaluations of teaching. *Econ. Lett.* **2019**, *179*, 66–71. [[CrossRef](#)]
19. Fan, Y.; Shepherd, L.J.; Slavich, D.; Waters, D.; Stone, M.; Abel, R.; Johnston, E.L. Gender and cultural bias in student evaluations: Why representation matters. *PLoS ONE* **2019**, *14*, 0209749. [[CrossRef](#)]
20. Chávez, K.; Mitchell, K.M.W. Exploring bias in student evaluations: Gender, race, and ethnicity. *Political Sci. Politics* **2020**, *53*, 270–274. [[CrossRef](#)]
21. Wang, L.; Gonzalez, J.A. Racial/ethnic and national origin bias in SET. *Int. J. Organ. Anal.* **2020**, *28*, 843–855. [[CrossRef](#)]
22. Kreitzer, R.J.; Sweet-Cushman, J. Evaluating student evaluations of teaching: A review of measurement and equity bias in SETs and recommendations for ethical reform. *J. Acad. Ethics* **2022**, *20*, 73–84.
23. Basow, S.; Codos, S.; Martin, J. The effects of professors' race and gender on student evaluations and performance. *Coll. Stud. J.* **2013**, *47*, 352–363.
24. Heffernan, T. Abusive comments in student evaluations of courses and teaching: The attacks women and marginalized academics endure. *High. Educ.* **2023**, *85*, 225–239.
25. Jones, J.; Gaffney-Rhys, R.; Jones, E. Handle with care! An exploration of the potential risks associated with the publication and summative usage of student evaluation of teaching (SET) results. *J. Furth. High. Educ.* **2014**, *38*, 37–56.
26. Tucker, B. Student evaluation surveys: Anonymous comments that offend or are unprofessional. *High. Educ.* **2014**, *68*, 347–358. [[CrossRef](#)]
27. Uttl, B.; Smibert, D. Student evaluations of teaching: Teaching quantitative courses can be hazardous to one's career. *PeerJ* **2017**, *5*, e3299. [[CrossRef](#)]
28. DiPietro, M.; Faye, A. Online student-ratings-of-instruction (SRI) mechanisms for maximal feedback to instructors. In Proceedings of the 30th Annual Meetings of the Professional and Organizational Development Network, Milwaukee, WI, USA, 2005.
29. Hamermesh, D.S.; Parker, A. Beauty in the classroom: Instructors' pulchritude and putative pedagogical productivity. *Econ. Educ. Rev.* **2005**, *24*, 369–376. [[CrossRef](#)]
30. Krautmann, A.C.; Sander, W. Grades and student evaluations of teachers. *Econ. Educ. Rev.* **1999**, *18*, 59–63. [[CrossRef](#)]
31. Matos-Díaz, H.; Ragan, J.F., Jr. Do student evaluations of teaching depend on the distribution of expected grade? *Educ. Econ.* **2010**, *18*, 317–330. [[CrossRef](#)]
32. Alauddin, M.; Kifle, T. Does the student evaluation of teaching instrument really measure instructors' teaching effectiveness? An econometric analysis of students' perceptions in economics courses. *Econ. Anal. Policy* **2014**, *44*, 156–168.
33. Bacon, D.R.; Johnson, C.J.; Stewart, K.A. Nonresponse bias in student evaluations of teaching. *Mark. Educ. Rev.* **2016**, *26*, 93–104.
34. Layne, B.H.; Decristoforo, J.R.; McGinty, D. Electronic versus traditional student ratings of instruction. *Res. High. Educ.* **1999**, *40*, 221–232. [[CrossRef](#)]
35. Avery, R.J.; Bryant, W.K.; Mathios, A.; Kang, H.; Bell, D. Electronic course evaluations: Does an online delivery system influence student evaluations? *J. Econ. Educ.* **2006**, *37*, 21–37. [[CrossRef](#)]
36. Nowell, C.; Gale, L.R.; Kerkvliet, J. Non-response bias in student evaluations of teaching. *Int. Rev. Econ. Educ.* **2014**, *17*, 30–38.
37. Heckman, J.J. Sample selection bias as a specification error. *Econometrica* **1979**, *47*, 153–161.
38. Bleske-Rechek, A.; Michels, K. RateMyProfessors.com: Testing assumptions about student use and misuse. *Pract. Assess. Res. Eval.* **2010**, *15*, 5.
39. Coladarci, T.; Kornfield, I. RateMyProfessors.com versus formal in-class student evaluations of teaching. *Pract. Assess. Res. Eval.* **2019**, *12*, 6.
40. Timmerman, T. On the validity of RateMyProfessors.com. *J. Educ. Bus.* **2008**, *84*, 55–61.
41. Albrecht, W.S.; Hoopes, J. An empirical assessment of commercial web-based professor evaluation services. *J. Account. Educ.* **2009**, *27*, 125–132. [[CrossRef](#)]
42. Brown, M.J.; Baillie, M.; Fraser, S. Rating Ratemyprofessors.com: A Comparison of Online and Official Student Evaluations of Teaching. *Coll. Teach.* **2009**, *57*, 89–92.
43. Sonntag, M.E.; Bassett, J.F.; Synder, T. An empirical test of the validity of student evaluations of teaching made on RateMyProfessors.com. *Assess. Eval. High. Educ.* **2009**, *34*, 499–504. [[CrossRef](#)]
44. Otto, J.; Sanford, D.A., Jr.; Ross, D.N. Does ratemyprofessor.com really rate my professor? *Assess. Eval. High. Educ.* **2008**, *33*, 355–368.
45. Hartman, K.B.; Hunt, J.B. What Ratemyprofessors.com reveals about how and why students evaluate their professors: A glimpse into the student mind-set. *Mark. Educ. Rev.* **2013**, *23*, 151–162.

46. Chou, S.Y.; Luo, J.; Ramser, C. High-quality vs low-quality teaching: A text-mining study to understand student sentiments in public online teaching reviews. *J. Int. Educ. Bus.* **2021**, *14*, 93–108. [[CrossRef](#)]
47. Constand, R.L.; Pace, R.D. Student evaluations of finance faculty: Perceived difficulty means lower faculty evaluations. *J. Financ. Educ.* **2014**, *40*, 14–28, 30–44.
48. Boehmer, D.M.; Wood, W.C. Student vs. faculty perspectives on quality instruction: Gender bias, 'hotness,' and 'easiness' in evaluating teaching. *J. Educ. Bus.* **2017**, *92*, 173–178.
49. Constand, R.L.; Pace, R.D.; Clarke, N. Accounting faculty teaching ratings: Are they lower because accounting classes are more difficult? *J. Account. Financ.* **2016**, *16*, 70–86.
50. Constand, R.L.; Clarke, N. How class and professor characteristics are related to finance faculty teaching ratings. *J. Financ. Educ.* **2017**, *43*, 101–119.
51. Constand, R.L.; Clarke, N.; Morgan, M. An analysis of the relationships between management faculty teaching ratings and characteristics of the classes they teach. *Int. J. Manag. Educ.* **2018**, *16*, 166–179.
52. Carter, R.E. Faculty scholarship has a profound positive association with student evaluations of teaching—Except when it doesn't. *J. Mark. Educ.* **2016**, *38*, 18–36. [[CrossRef](#)]
53. Smith, K. Getting econometrics students to evaluate student evaluations. In *Shaping the Learning Curve: Essays on Economic Education*; Mixon, F.G., Jr., Ed.; iUniverse: New York, NY, USA, 2005; pp. 15–25.
54. Mixon, F.G., Jr.; Smith, K.W. Instructor attractiveness and academic rigour: Examination of student evaluation data. *Australas. J. Econ. Educ.* **2013**, *10*, 1–13.
55. Green, T.G.; Mixon, F.G., Jr.; Treviño, L.G. Have you seen the new econ prof? Beauty, teaching, and occupational choice in higher education. In *Shaping the Learning Curve: Essays on Economic Education*; Mixon, F.G., Jr., Ed.; iUniverse: New York, NY, USA, 2005; pp. 57–67.
56. Green, T.G.; Mixon, F.G., Jr.; Treviño, L.J. Instructor attractiveness and institutional choice in economics: A decomposition approach. In *New Developments in Economic Education*; Mixon, F.G., Jr., Cebula, R.J., Eds.; Edward Elgar Publishing: Northampton, MA, USA, 2014; pp. 209–217.
57. Osoian, C.; Nistor, R.; Zaharie, M.; Flueraș, H. Improving higher education through student satisfaction surveys. In Proceedings of the 2nd International Conference on Education Technology and Computer, Shanghai, China, 22–24 June 2010; Volume 2, pp. 436–440.
58. Benton, S.L.; Cashin, W.E. *Student Ratings of Teaching: A Summary of Research and Literature*; IDEA Paper Series; ERIC: Washington, DC, USA, 2014.
59. Sabatier, M. Do female researchers face a glass ceiling in France? A hazard model of promotions. *Appl. Econ.* **2010**, *42*, 2053–2062. [[CrossRef](#)]
60. Cooray, A.; Verma, R.; Wright, L. Does a gender disparity exist in academic rank? Evidence from an Australian university. *Appl. Econ.* **2014**, *46*, 2441–2451.
61. Bukstein, D.; Gandelman, N. Glass ceilings in research: Evidence from a national program in Uruguay. *Res. Policy* **2019**, *48*, 1550–1563.
62. Thorndyke, L.E.; Milner, R.J.; Jaffe, L.A. Endowed chairs and professorships: A new frontier in gender equity. *Acad. Med.* **2022**, *97*, 1643–1649. [[CrossRef](#)]
63. White, H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* **1980**, *48*, 817–838. [[CrossRef](#)]
64. Rosenbaum, P.R.; Rubin, D.B. The central role of the propensity score in observational studies for causal effects. *Biometrika* **1983**, *70*, 41–55.
65. Narita, K.; Tena, J.D.; Detotto, C. Causal inference with observational data: A tutorial on propensity score analysis. *Leadersh. Q.* **2023**, *34*, 101678. [[CrossRef](#)]
66. Imbens, G.W. The role of the propensity score in estimating dose-response functions. *Biometrika* **2000**, *87*, 706–710. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.