Student Evaluations of Teaching: Does Pedagogy Matter?

Howard H. Cochran, Jr., Gary L. Hodgin, and Joachim Zietz *

Abstract

The object of this study is to determine the extent to which the pedagogical choice of an instructor may influence the student evaluation of teaching effectiveness in Principles of Economics. Estimates of an educational production function are made using a large national data set (TUCE III) available from the National Council on Economic Education. Significant gains in student satisfaction are brought about by intensive use of homework, comprehensive final exams, class discussions, and frequent feedback. Additionally, instructor enthusiasm, class preparation, communication style, and grading standards contribute to the formation of student perceptions.

Introduction

Student evaluations of teaching (SETs) are widely used as a measure of teaching effectiveness in higher education across the United States. In a survey of 114 economics departments in 1992, White (1995, 81-82) found that SETs were by far the most common instrument for measuring teaching effectiveness with 75 of 76 respondents using SETs and the other planning to use them in the future. In contrast, only about 40 percent of the institutions in higher education used SETs in 1960 according to Stechlein (1960) while Wilson (1998) reports that only 30 percent of institutions asked students to evaluate professors in 1973. SETs are now the single most important ingredient in formal assessment strategies imposed on faculty.[1] This is not surprising, as accreditation agencies require the evaluation and documentation of instructional effectiveness.[2]

Braskamp and Ory (1994) are rather typical in their belief that SETs are valid measures of instructor effectiveness when they are administered and interpreted properly. Proper interpretation, however, requires an understanding of what causes SET scores to vary and by how much. Over the

* Howard H. Cochran, Jr., cochranh@mail.belmont.edu, and Gary L. Hodgin, hodging@mail.belmont.edu, Department of Economics and Finance, Belmont University, 1900 Belmont Boulevard, Nashville, TN 37212; and Joachim Zietz, Department of Economics and Finance, Middle Tennessee State University, Murfreesboro, TN 37132, jzietz@mtsu.edu.

[1] Becker and Watts (1999) find that SETs dominate the evaluation of teaching although institutions are exploring other methods such as peer evaluation, classroom observation, grade distribution, and longitudinal student surveys.

[2] For example, the AACSB standards for accreditation (AACSB International, 1999, FD.3.b, page 12) require a formal process to evaluate teaching effectiveness. Similarly, the SACS criteria for accreditation (Southern Association of Colleges and Schools, 1996, 4.2.4 page 31) mandate the survey of student opinion to evaluate instructional programs.

past 30 years, there have been numerous empirical studies on the determinants of SET scores.[3] However, one shortcoming of most of these studies is that little or nothing is said about how the variation in SET scores is attributed to pedagogy.

A few studies have empirically identified how pedagogical choices made by the teacher influence SET scores. Leeds et al. (1998) find that only two pedagogical variables have a significant impact on SET scores, the traditional lecture method and stopping the lecture to check for understanding. Both variables have a positive influence on SET scores for college economics classes. Morgan and Vasche (1978) and Bejar and Doyle (1978) find traditional lectures and class discussions to have a positive impact as well. Using a more comprehensive group of pedagogy variables, Boex (2000) finds that presentation ability, organization and clarity, grading, and assignments have positive influences on SET scores.[4] However, the extent to which the above findings can be generalized is questionable because these studies are institutionally specific and use class averages rather than individual student data. With the exception of Boex (2000), only carefully selected pedagogical variables are typically analyzed. These limitations may be one factor contributing to faculty apprehensions about using SETs in evaluating faculty (Gomez-Mejia and Balkin 1992).

This study uses a richer data set than previous studies to analyze the impact of pedagogy on SET scores. A large cross-sectional data set covering many different institutions is employed along with individual student and instructor data rather than class averages. This study also contributes to the literature on the scholarship of teaching economics by testing the effectiveness of several pedagogical choices. Some of these choices are unique to this study, and the results obtained can be more easily generalized than from previous inquiries.

This paper is organized as follows. The following section discusses the data set and the methodological approach. Next, the estimation results are discussed and compared to those found in the literature. The paper ends with a summary of main points and suggestions for further study.

Data and Methodology

The TUCE III data set (Saunders 1994) includes SET scores from 93 Principles of Macroeconomic classes and 96 Principles of Microeconomics classes. The evaluations were administered at the end of the semester. Fifty-three separate institutions from four Carnegie classifications are represented within the data. The total data set contains 9,679 observations with information compiled from faculty questionnaires, student questionnaires, and student test scores along with the student evaluations of teaching. Each observation represents an individual student. Since answering the questionnaires was voluntary, missing values exist for some students. Eliminating all observations with missing values for the selected variables leads to a data set of 3,322 observations.

The independent variables (Table 1) are organized into two broad groups: pedagogy and non-pedagogy variables. The non-pedagogy variables can be separated into several distinct groups: (a) variables describing the institutional characteristics in general terms, such as private versus public institutions; (b) variables on class characteristics not under the control of the instructor, such as class meetings per week; (c) instructor characteristics, such as years of teaching experience, and; (d) student characteristics, such as grade point average.

---

[3] See Brashkamp and Ory (1994, 173-191) for a survey of the literature in education. In addition, there are several specific studies in economics, mostly from the introductory courses. Becker (1997, 1368-1369) and Boex (2000, 211-227) provide a brief summary of this literature.

[4] The terms "instructional effectiveness" and "student satisfaction" are used interchangeably throughout this paper. Our objective is to investigate the relationship between SET scores and instructors' choices of pedagogy. Whether students' responses to questions about pedagogy reflect satisfaction or teaching effectiveness is not addressed in this paper.

TABLE 1. VARIABLE DEFINITIONS AND BASIC STATISTICS (3,322 OBSERVATIONS, EXCEPT WHERE NOTED)

Variable	Definition	Mean	Standard Deviation	Min.	Max.
Dependent Variable	Student Evaluation of Teaching Effectiveness	40.364	8.47	3	50
Doctoral Institut.	1 = Doctoral Institution	0.254	0.44	0	1
Comprehen. Inst.	1 = Comprehensive Institution	0.478	0.50	0	1
Liberal Arts Inst.	1 = Liberal Arts Institution	0.122	0.33	0	1
Private Inst.	1 = Private Institution	0.207	0.41	0	1
Macroeconomics	1 = Macroeconomics class	0.478	0.50	0	1
Class per week	Class Meetings per Week	2.723	0.65	1	5
Class size	Number of students in class	70.508	44.05	10	232
Native English	1 = English Is Native Language of Instructor	0.965	0.18	0	1
Regular faculty	1 = Regular Faculty	0.745	0.44	0	1
Adjunct faculty	1 = Adjunct Faculty	0.063	0.24	0	1
Grad. stud. fac.	1 = Graduate Student Faculty	0.091	0.29	0	1
Tenured		0.678	0.47	0	1
Course per term	Instructor Courses per Term	2.034	0.79	1	4
Article published	1 = Article Published in Last Five Years	0.336	0.47	0	1
Doctoral Degree	1 = Highest Degree Earned Is a Doctorate	0.586	0.49	0	1
Years teaching	Number of Years Teaching	13.933	8.86	1	34
Yrs teach course	Number of Years Teaching Course	11.562	8.72	1	32
Male[1]	1 = Student Gender Is Male	0.549	0.50	0	1
GPA[2]	Student Grade Point Average	2.950	0.66	0	6
College calculus	1 = Student Had College Calculus	0.464	0.50	0	1
College econ.	1 = Student Had College Economics	0.529	0.50	0	1
High school econ.[3]	1 = Student Had High School Economics	0.477	0.53	0	2
Load part time	1 = Student Course Load Is Part-Time	0.126	0.33	0	1
Load above avg.	1 = Student Course Load Is Above Average	0.131	0.34	0	1
Job	1 = Student Has a Job	0.561	0.50	0	1
Pctteach	Percent of Time in Teaching	62.440	23.56	5	99
Pctrsch	Percent of Time in Research	19.656	20.34	0	80
Pctadm	Percent of Time in Administration	10.371	15.46	0	95
Pctcons	Percent of Time in Consulting	1.931	5.51	0	50
Sametext	1 = Same Text used as Last Year	0.722	0.45	0	1
Stdguide	1 = Study Guide Used Intensively	0.191	0.39	0	1
Homewrk	1 = Home Work Used Intensively	0.330	0.47	0	1
Readhand	1 = Reading and Handouts Used Intensively	0.143	0.35	0	1
Computer	1 = Computers Used Intensively	0.029	0.17	0	1
Videos	1 = Videos Used Intensively	0.043	0.20	0	1
Pctgrdqz	Percent of Grade for Quizzes	11.023	14.94	0	72
Pctgrdhom	Percent of Grade for Home Work	3.662	7.00	0	40
Pctgrdmt	Percent of Grade for Mid-Terms	55.144	17.98	0	99
Pctgrdpap	Percent of Grade for Papers	0.614	2.73	0	20
Pctfincomp	Percent of Final Exam that is Comprehensive	58.458	41.73	0	99
Pcttffingrd	True-False Percent as Percent of Final Grade	4.548	10.45	0	60
Pctmcfingrd	Multiple Choice Percent of Final Grade	73.261	27.92	0	99
Pctessay	Written Answers as Percent of Final Grade	19.451	25.46	0	100
Pctlect	Percent of Time Spent on Lecturing	77.641	12.71	33	99
Pctdialog	Percent Spent on Instructor-Student Dialogue	21.493	13.02	0	66
Expgrade[4]	Student Expected Course Grade	28.245	8.09	0	54

Notes: [1] 3299 obs., [2] 2856 obs., [3] 3314 obs., [4] 3231 obs.

The estimation approach follows a production function framework similar to Hanushek (1979), Manahan (1983), and Becker and Walstad (1987). Our application of this approach can be summarized as follows. First, student evaluations of teaching are regressed on a set of non-pedagogy variables. This set of variables is selected from all available variables in the TUCE III on the basis of (a) the findings of earlier studies and (b) the effort to keep the number of observations reasonably large.[5] Hence, variables that are typically not statistically significant or have very few observations are not included in the set of independent variables. Next, we test to what extent this base model can be simplified by dropping non-significant variables.

Second, a full complement of pedagogy variables is added to the reduced variable set that is left after dropping insignificant variables from the base model. Again, insignificant variables are dropped to increase the efficiency of the statistical estimates.

Third, we examine to what extent the regression results are sensitive to the inclusion or exclusion of the variable "expected grade." Previous studies (Boex 2000; Greenwald and Gillmore 1997; Mason, Steagall, and Fabritius 1995; Nelson and Lynch 1984) find that expected grade exerts a positive influence on SET scores. However, the controversy surrounding this finding is in its interpretation. What is at issue is whether or not instructors are buying higher evaluations by awarding higher grades, or if more effective instructors receive higher evaluations because their students learn more, or if learning and satisfaction result form high student motivation (Cashin 1995). Thus, "expected grade" is added to the reduced variable set from the second step in order to avoid misspecification of the model.

Fourth and finally, SET scores are regressed separately on student and on instructor evaluations of four broad categories of instructor behavior: enthusiasm, class preparation, communication, and grading. The regression results allow one to examine whether instructors and students have broadly similar views on what is important for overall teaching excellence as it is manifested in student evaluations.

Estimation Results and Interpretation

The estimation results are provided in Tables 2 through 4. Table 2 reports the results of regressions in which SET is regressed on non-pedagogy variables. The first three variables in Table 2 are institutional variables. Since the base institutional category is the two-year college, the coefficients for doctoral, comprehensive, and liberal arts institutions measure the difference between these institutions and the two-year college. The results indicate that instructors at doctoral and comprehensive institutions receive significantly lower student evaluations. Perhaps students at these institutions feel a greater distance between themselves and their instructors and consequently evaluate them more harshly. There appears to be no difference in SET scores between private and public institutions; nor does there appear to be any difference between liberal arts and two-year institutions.

Each of the course variables is statistically significant. Since "microeconomics" is the base variable in the regression equation, the negative coefficient on macroeconomics indicates that macroeconomics instructors receive lower evaluations than microeconomics instructors. Similarly, instructors of larger classes and classes meeting less frequently score higher on the student evaluations. Interestingly, institution size and class size have offsetting influences on SET scores. The higher scores in larger classes may result from more effective teachers being assigned to these sections, a finding that is consistent with Boex (2000). Alternatively, it may be that lectures are the preferred learning style for principles of economics and are predominantly found in large classes.

[5] The danger of small data sets is that the estimates pick up the characteristics of individual classes that may play a significant role in smaller data sets but that may not be very representative. Large swings in coefficient estimates and even signs can easily occur in small samples.

While Braskamp and Ory (1994) cite studies that contradict this finding, none of these studies discuss the preferred learning method of students in the classes that are evaluated nor do they use regression analysis to determine the individual influence of variables. Similar to Leeds et al. (1998), there is a negative relationship between SET scores and the number of class meetings per week.

TABLE 2. REGRESSIONS WITH ALL BUT PEDAGOGY VARIABLES

	All variables included			Zero restrictions imposed for the same observations			Zero restrictions imposed for more observations		
Variables	coefficient	t-value	p-value	coefficient	t-value	p-value	coefficient	t-value	p-value
Constant	43.404	36.32	0.00	43.403	43.04	0.00	43.807	47.44	0.00
Doctoral Institut.	-1.498	-2.21	0.03	-1.766	-2.76	0.01	-1.852	-3.20	0.00
Comprehen. Inst.	-1.328	-2.37	0.02	-1.447	-2.61	0.01	-1.535	-3.08	0.00
Liberal Arts Inst.	0.439	0.62	0.54	0.065	0.10	0.92	0.107	0.18	0.85
Private Inst.	-0.470	-1.19	0.24						
Macroeconomics	-0.816	-2.83	0.01	-0.828	-2.89	0.00	-0.908	-3.48	0.00
Class per week	-0.692	-2.97	0.00	-0.699	-3.00	0.00	-0.782	-3.62	0.00
Class size	0.011	2.92	0.00	0.011	3.16	0.00	0.013	3.93	0.00
Native English	4.430	7.44	0.00	4.412	7.47	0.00	3.856	7.52	0.00
Regular faculty	1.824	2.84	0.01	1.829	2.86	0.00	1.938	3.40	0.00
Adjunct faculty	-0.592	-0.91	0.36	-0.551	-0.85	0.40	-0.597	-1.00	0.32
Grad. stud. fac.	-2.239	-2.66	0.01	-1.952	-2.44	0.02	-1.344	-1.90	0.06
Tenured	-2.521	-4.08	0.00	-2.529	-4.12	0.00	-2.673	-4.86	0.00
Course per term	-1.292	-5.51	0.00	-1.327	-5.72	0.00	-1.177	-5.63	0.00
Article published	0.909	2.10	0.04	0.910	2.10	0.04	0.742	1.88	0.06
Doctoral Degree	-1.263	-2.74	0.01	-1.212	-2.65	0.01	-1.120	-2.76	0.01
Years teaching	-0.095	-2.72	0.01	-0.096	-2.80	0.01	-0.077	-2.40	0.02
Yrs teach course	0.053	1.42	0.16	0.058	1.60	0.11	0.054	1.62	0.11
Male	0.237	0.86	0.39						
GPA	-0.071	-0.34	0.74						
College calculus	-0.195	-0.65	0.51	-0.259	-0.90	0.37	-0.380	-1.43	0.15
College econ.	0.425	1.40	0.16	0.433	1.43	0.15	0.367	1.34	0.18
High school econ	0.111	0.44	0.66						
Load part time	-1.019	-2.22	0.03	-0.998	-2.20	0.03	-1.026	-2.49	0.01
Load above avg.	-0.236	-0.58	0.56	-0.237	-0.59	0.56	-0.304	-0.82	0.41
Job	-0.469	-1.67	0.10	-0.445	-1.59	0.11	-0.579	-2.23	0.03
Number of obs.	3982			3982			4627		
R^2	0.0451			0.0446			0.0436		

Notes: See Table 1 for variable definitions.

Instructor-specific variables that positively contribute to student satisfaction include English as the first language of the professor, being a full-time faculty member, and engaging in research. English as a first language may allow for better communication between instructors and students, a finding consistent with that of Leeds et al. (1998) and Anderson and Seigfried (1997). However, the relative size of English as a first language is reduced when pedagogical choices are included. A full-time regular faculty member presumably allows students to have better access to the professor, which raises student satisfaction. Engaging in research may make instructors more active in their discipline and as a result more knowledgeable and interesting in the classroom.

Tenure tends to reduce student evaluations. Perhaps the tenure process at many universities underemphasizes the importance of teaching effectiveness or post-tenure reviews may not provide adequate incentives to maintain teaching effectiveness. Tenured faculty may tend to rest on their laurels and may not be willing to go the extra mile in their teaching endeavors. As Johnston, McDonald, and Williams (2001) suggest, incentive structures encouraging faculty to devote more time to teaching improvement have been historically absent. However, Becker and Watts (1999) report a recent trend to weight personnel decisions regarding tenure, promotion, and salary increases more heavily in favor of teaching performance. Thus, a realignment of the incentive structure may eventually nudge tenured faculty to enhance future teaching performance (Gomez-Meija and Balkin 1992).

Why does a doctoral degree reduce evaluations? Faculty members that hold doctoral degrees may prefer to teach courses in their area of specialization. For them, teaching principles of economics places them at a comparative disadvantage. As Anderson and Siegfried (1997) found with respect to Principles of Macroeconomics, we also find a negative relationship between "years teaching" and SET (Feldman 1983). This may indicate that instructors with a long teaching career may be less responsive to students because of burnout or worn-down sensitivities. Additionally, as the number of course preparations increases per semester, student satisfaction declines. Faculty members having to allocate their time among several preparations are probably less likely to deliver any one course as well. By contrast, fewer preparations taught multiple times allows for continuous improvement during the term. A similar argument explains why SET scores are positively related to the variable "years teaching course."

There is nothing remarkable to report with regard to the student variables. Only one variable (part-time course load) is statistically significant. Grade point average is negative but not significant, a finding confirmed, inter alia, by Leeds et al. (1998) and Feldman (1976). Gender does not influence the overall evaluation of instructors, which confirms the findings of Basow and Silberg (1987) along with Feldman (1993). Part-time students appear to give lower student evaluations. This may be because the opportunity cost of their time may be higher than for full-time students. They cannot devote enough time to studying to fully understand what is going on and, as a consequence, they take their frustration out on the instructor. A similar argument applies to students with a regular job. Needham (1978) argues that students allocate their time to course work in an attempt to maximize their grade point average. The more time instructors require of students, the less satisfaction the student receives from the course because of the higher opportunity costs for a given grade, a prediction consistent with utility maximizing models of student achievement versus leisure (McKenzie and Staaf 1974; Kelley 1975). If studying economics takes more time and effort, satisfaction will decline.

One should note that there is little variation in the estimated coefficients reported in Table 2 as one restricts some variable coefficients to zero. The restrictions are supported by simple F-tests at very high levels of statistical significance. Eliminating variables from the regression equation means that the potential number of observations goes up because observations containing missing values were omitted from the initial regression. In the case of the regression of Table 2, removing the variables *Male* and *GPA* adds about 650 observations. However, the additional observations have little effect on the estimated coefficients or their statistical significance. Consequently, the explanatory power of the model is maintained even with fewer independent variables.[6] The restricted model of Table 2 with 4,627 observations becomes the base model for Table 3.

---

[6] We note that the relatively low R^2 in our study is similar to other studies within the same genre. While the results of our study are not deterministic, they are indicative of tendencies toward effective pedagogical clues. Clearly, more research could potentially enhance our understanding of variations in student evaluation scores.

TABLE 3. REGRESSIONS WITH ALL VARIABLES INCLUDING PEDAGOGY VARIABLES

	All variables included			Zero restrictions imposed for same observations			Expected grade added to restricted model			Without expected grade for same observations		
Variables	coef.	t-value	p-value	coef.	t-value	p-value	coef.	t-value	p-value	coef.	t-value	p-value
Constant	53.052	8.39	0.00	52.331	24.96	0.00	45.478	20.80	0.00	51.967	24.68	0.00
Doctoral Inst.	-2.191	-2.30	0.02	-2.440	-2.92	0.00	-2.278	-2.74	0.01	-2.627	-3.11	0.00
Comprehen. Inst.	-1.785	-2.44	0.02	-1.951	-2.85	0.00	-1.813	-2.67	0.01	-2.066	-2.99	0.00
Liberal Arts Inst.	-0.286	-0.27	0.78	-0.435	-0.53	0.60	-0.316	-0.38	0.70	-0.478	-0.57	0.57
Macroeconomics	0.431	1.22	0.22	0.344	1.01	0.31	0.643	1.90	0.06	0.497	1.44	0.15
Class per week	-0.697	-2.13	0.03	-0.645	-2.33	0.02	-0.515	-1.84	0.07	-0.680	-2.42	0.02
Class size	0.018	4.06	0.00	0.020	4.76	0.00	0.022	5.29	0.00	0.020	4.91	0.00
Native English	1.754	1.79	0.07	1.767	1.99	0.05	1.927	2.20	0.03	1.900	2.12	0.03
Regular faculty	1.644	1.39	0.17	1.284	1.33	0.18	1.293	1.35	0.18	1.246	1.27	0.20
Adjunct faculty	-2.398	-2.22	0.03	-2.196	-2.46	0.01	-1.795	-1.99	0.05	-2.228	-2.46	0.01
Grad. stud. fac.	-3.513	-3.25	0.00	-3.776	-4.00	0.00	-3.980	-4.20	0.00	-3.799	-3.97	0.00
Tenured	-2.764	-2.48	0.01	-2.456	-2.54	0.01	-1.690	-1.76	0.08	-2.321	-2.37	0.02
Course per term	-1.639	-5.73	0.00	-1.820	-7.24	0.00	-1.977	-7.77	0.00	-1.883	-7.38	0.00
Article published	1.574	2.45	0.01	1.698	2.93	0.00	1.539	2.62	0.01	1.532	2.58	0.01
Doctoral Degree	-1.991	-2.98	0.00	-1.728	-3.05	0.00	-1.603	-2.84	0.00	-1.593	-2.78	0.01
Years teaching	-0.017	-0.32	0.75	-0.033	-0.67	0.51	-0.032	-0.64	0.52	-0.039	-0.77	0.44
Yrs. teach course	0.047	0.94	0.35	0.058	1.20	0.23	0.058	1.21	0.23	0.064	1.29	0.20
College calculus	-0.522	-1.63	0.10	-0.515	-1.64	0.10	-1.136	-3.57	0.00	-0.501	-1.58	0.12
College econ.	0.014	0.04	0.97	-0.021	-0.06	0.95	-0.067	-0.20	0.84	-0.018	-0.05	0.96
Load part time	-0.903	-1.85	0.06	-0.948	-1.97	0.05	-0.768	-1.61	0.11	-0.762	-1.56	0.12
Load above avg.	-0.635	-1.43	0.15	-0.624	-1.41	0.16	-0.818	-1.84	0.07	-0.613	-1.36	0.18
Job	-0.540	-1.78	0.08	-0.547	-1.81	0.07	-0.486	-1.60	0.11	-0.585	-1.90	0.06
Pctteach	-0.005	-0.34	0.74									
Pctrsch	0.000	0.02	0.98									
Pctadm	0.014	0.85	0.40									
Pctcons	-0.089	-2.33	0.02	-0.093	-2.66	0.01	-0.089	-2.49	0.01	-0.090	-2.52	0.01
Sametext	-1.833	-3.71	0.00	-2.060	-4.83	0.00	-2.127	-5.07	0.00	-2.165	-5.02	0.00
Stdguide	0.542	0.91	0.36	0.427	0.78	0.44	0.547	1.00	0.32	0.402	0.72	0.47
Homewrk	1.765	3.86	0.00	1.732	3.97	0.00	1.894	4.39	0.00	1.838	4.18	0.00
Readhand	0.881	1.60	0.11	0.767	1.50	0.13	0.459	0.89	0.37	0.565	1.09	0.28
Computer	0.928	0.54	0.59	0.807	0.49	0.62	1.540	0.95	0.34	1.297	0.77	0.44
Videos	-1.941	-1.90	0.06	-1.721	-1.82	0.07	-2.924	-3.10	0.00	-1.998	-2.07	0.04
Pctgrdqz	-0.107	-3.65	0.00	-0.110	-4.09	0.00	-0.102	-3.85	0.00	-0.104	-3.83	0.00
Pctgrdhom	-0.122	-3.19	0.00	-0.127	-4.19	0.00	-0.124	-4.13	0.00	-0.124	-4.07	0.00
Pctgrdmt	-0.099	-4.07	0.00	-0.104	-4.60	0.00	-0.098	-4.42	0.00	-0.098	-4.27	0.00
Pctgrdpap	-0.035	-0.41	0.68	-0.034	-0.46	0.64	-0.029	-0.40	0.69	-0.012	-0.17	0.87
Pctfincomp	0.011	2.02	0.04	0.013	2.60	0.01	0.014	2.82	0.01	0.014	2.77	0.01
Pcttffingrd	0.008	0.29	0.77									
Pctmcfingrd	0.007	0.39	0.69									
Pctessay	0.007	0.38	0.70									
Pctlect	-0.019	-0.34	0.73									
Pctdialog	0.055	1.08	0.28	0.078	5.36	0.00	0.079	5.44	0.00	0.074	5.04	0.00
Expgrade							0.194	9.98	0.00			
Number of obs.	3322			3322			3231			3231		
R^2	0.0948			0.0942			0.1267			0.0960		

Notes: See Table 1 for variable definitions and basic statistics.

Table 3 gets to the heart of the paper as the pedagogy variables are added to the variables examined in Table 2. In particular, the first regression of Table 3 includes the variables from the restricted regression of Table 2 and all pedagogy variables that are considered. It is apparent that these variables, taken jointly, have significant explanatory power. The coefficient of determination (adjusted R^2) more than doubles with their inclusion. The available number of observations declines somewhat due to missing values for some of the new variables. Both the change in the number of variables and the change in the set of observations affects some of the variables that are examined by the regressions of Table 2. The most dramatic change occurs for the variable "macroeconomics." Its sign reverses from negative to positive and is no longer statistically significant, a finding that is consistent with Leeds et al. (1998). Hence, if one accounts for pedagogy, then macroeconomics classes are not less liked by students than microeconomics classes. Why macroeconomics classes reduce SET scores in Table 2 but not in Table 3 can be understood if one considers the following: Siegfried et al. (1996) report that microeconomic courses heavily use homework and macroeconomic courses are more likely to use videos. Homework positively contributes to course satisfaction, while videos reduce student affinity. The negative sign for macroeconomics in Table 2 therefore reflects the influence of a pedagogy that students prefer and also one that they do not particularly like.

Some variables that create positive feelings among students are reduced in magnitude in Table 3 relative to Table 2. Among these variables is the English language variable, the fact that the instructor is regular faculty as opposed to "other," the time the instructor has been teaching the course, or whether students have taken previous college economics courses. Academic research demonstrates a stronger contribution to SET when accounting for pedagogy, a noteworthy difference from studies that do not account for instructional choices (Yunker and Marlin 1984; Feldman 1987).

Other variables turn out to exert more of a negative influence on student satisfaction than previously measured: institutional type, adjunct faculty status, tenure, courses taught per term, terminal qualifications, college calculus, and student allocations of time to either full-time employment or above-average course loads. Part-time enrollment has now less of a negative impact on SET scores. Years teaching experience and regular faculty status are now statistically insignificant, perhaps a result of more experienced instructors adopting more effective pedagogies. Class size remains relatively unchanged.

The added pedagogy variables contribute to SET in four dimensions: professor allocation of time, selection of pedagogical instruments, course organization, and the degree of in-class dialogue.

The first four variables of the pedagogy set relate to the way an instructor allocates time among competing uses. All other use of professorial time serves as the base variable. Teaching, research, administration, and consulting do not appear to contribute to student satisfaction as much as alternative uses of time such as mentoring individual students, giving personal attention, or providing frequent feedback on performance. The category "other" is the base category in this case. Hence, the four estimated coefficients provide the difference relative to the "other" category. Spending a greater percentage of time in teaching, research, or administrative endeavors does not contribute to improving SET (Yunker and Marlin 1984). As long as students have access to professors, students may not have ill feelings toward how a professor spends non-class time (Tang 1997). It is apparent that time spent on consulting affects teaching evaluations negatively. This may be because consulting reduces a professor's access to students or takes the instructor's mind off the teaching process more so than other activities. The insignificance of time allocated to research and the significance of having had an article published are not necessarily contradictory. What matters to students is a professor's contemporary engagement in the discipline, regardless of the amount of time that is required to reach that level of inquiry.

Some pedagogical instruments are more effective than others at providing greater course satisfaction. Intensive use of homework assignments yields significant gains while intensive use of non-text readings and handouts shows much less importance. Mehdizadeh (1990) attributes this to the value that students place on supplementary materials. The fact that the use of the same text as in the last semester affects teaching evaluations negatively is not easily explained. One possible explanation is that instructors tend to be less careful in their class preparation. Quite apparently, students notice that, as this variable is one of the most significant ones among the pedagogy variables. Aigner and Thum (1986) find that preparation and SET scores are positively related. Another explanation could be a negative halo effect that causes students to rate a course lower in the current term because of their dislike of a textbook they used in a previous semester.

Interestingly, students give videos a negative vote as a pedagogy device. This is not all that surprising given the fact that today's students tend to be oversaturated with TV. Also, most video programs tend to be rather simplistic and unresponsive to students' particular interests and questions. As no interaction is taking place, students tend to get bored and turn off. The use of computers does not show up to be significant in its impact on student evaluations. However, this is likely the result of the fact that computer use was still in its infancy at the time the TUCE data were collected. Not only did very few classes get exposed to computers in the classroom, but the computer material was in all likelihood rather primitive compared to today. Students find intensive use of study guides detracting, but as previously mentioned enjoy the intensive use of homework problems. These results are not inconsistent because study guides are standardized while homework is tailored by the professor to the class and is indicative of preparation. Customization allows for higher satisfaction and possibly higher grades since the material over which a student is tested is more likely to come directly from the professor rather than the publisher.

The elements of course organization are primarily concerned with how the final grade is determined by the instructor. The coefficients of the four variables following "videos" must be interpreted as differences relative to the base category, which is made up of the percentage by which the final grade is determined by a category identified as "other" that is primarily inclusive of final exams. It is apparent that students consider the grading issue to be very important for their student evaluations, a finding that is common in the literature (Boex 2000; Tang 1997). By identifying various methods of assigning grades, this paper adds to our understanding of the connection between specific grading practices and student satisfaction. Three of the four grading variables are highly significant statistically. The negative coefficients suggest that students do not like grades being determined by quizzes, mid-terms or homework assignments relative to finals. This may be because many students feel that methods are extrinsic motivators designed to monitor one's learning progress while the outcome of the learning process can only be determined at the end of the course through either exams or other written assignments. Extrinsic motivators, while inducing a desired behavior, are less satisfying than intrinsic motivators (Herzberg 1987).

Course organization should also consider how a final exam is structured with respect to the four variables following "Pctgrdpap." In general, students are indifferent to using either more multiple choice or more essay questions on a final exam. Perhaps the perceived fairness in the grading of a final exam is more important than the selection of a particular type of question used in the determination of a final grade. Students tend to approve of exams that are more comprehensive. This appears to be consistent with the preference for final exams as a means to arrive at the course grade. But it may also be the case that the best and best-liked teachers are using comprehensive final exams because comprehensive finals are known to raise student learning (Zietz and Cochran 1998). Consequently, student satisfaction improves with the perception that comprehensive finals are a better reflection of the gain in knowledge (Bosshardt and Watts 2001; Watts and Lynch 1989; Leeds et al. 1998).

The degree of instructor-student dialogue in a course appears to be inconsequential. Both the percentage of class time spent in lecture and the percentage of class time spent in either students responding to instructor questions or the instructor responding to student questions will not improve teaching evaluations. Lecture and dialogue appear to be no more satisfying to students than other communication techniques. Perhaps this may shed light on the lack of enthusiasm for adopting active learning pedagogies along with the continuing predominance of the "chalk and talk" method of instruction (Becker and Watts 2001).

The second regression reported in Table 3 eliminates a number of statistically insignificant variables while forcing the set of observations to be identical to the one used for the first regression. There is no significant change in any of the coefficients or in the R^2, which is what one would expect from the deletion of jointly insignificant variables.

The third regression reported in Table 3 adds the variable "expected grade" to the reduced variable set of the second regression. Expected grade is a highly significant variable in all previous work on student evaluations and also turns out highly significant for this data set (Gramlich and Greenlee 1993; Aigner and Thum 1986; Stratton et al. 1994). The statistical significance of "expected grade" translates into a much higher R^2. As previously noted, the interpretation of coefficient on "expected grade" is ambiguous and goes beyond the scope of this analysis. Few variables are much affected by the inclusion of "expected grade." The variable for tenure is somewhat less significant with "expected grade" included in the regression, and the variables "college calculus," "pctdialog," and "videos" gain in statistical significance although they do not change in sign. In sum, this study confirms the finding of previous research that higher expected grades positively contribute to course evaluations. Table 3 adds to this the knowledge that the contribution of pedagogy to course satisfaction is largely unaffected by student grade expectations.

TABLE 4. PERCEIVED DETERMINANTS OF OVERALL TEACHING EVALUATION

	Instructor perceptions			Student perceptions		
Variable	coefficient	t-value	p-value	coefficient	t-value	p-value
Constant	2.390	4.90	0.00	-5.488	-7.58	0.00
Enthusiasm	0.440	47.91	0.00	0.413	22.22	0.00
Preparation	0.182	25.82	0.00	0.375	16.98	0.00
Communication	0.216	24.73	0.00	0.121	6.80	0.00
Grading	0.063	6.57	0.00	0.168	9.47	0.00
Number of obs.	5824			5824		
R^2	0.5555			0.5286		

Notes: For instructor perceptions, instructor=s own overall evaluation of teaching effectiveness is regressed on instructor=s own evaluations of enthusiasm, preparation, etc. For student perceptions, students' overall evaluation of teaching is regressed on students' evaluation of instructor enthusiasm, preparation, etc.

The first regression reported in Table 4 explains self-assessment of teaching in terms of the instructors' self-assessment of enthusiasm, class preparation, communication ability, and grading. These categories of assessment are significant determinants of student satisfaction (Lumsden 1974; Lumsden and Scott 1983; Boex 2000; Bosshardt and Watts 2001). The second reported regression does the same but for student evaluations of all categories. For both instructors and students, overall teaching evaluations are mainly driven by the enthusiasm that the instructor has for the subject matter. Whereas communication ability is ranking second in the minds of instructors, instructor class preparation ranks higher, in fact much higher, for students. It is also obvious that grading is much more on students' minds than communication ability.

These results of Table 4 offer some interesting insight into the sign and significance of some variables in Table 3. According to Table 4, instructor preparation is paramount to receiving higher evaluations. Use of videos and the same text may be perceived as substitutes for preparation. By contrast, intensive use of homework reflects greater preparation and as such improves satisfaction. Similarly, adjunct faculty and graduate student faculty may spend less time in preparation along with those faculty members who devote more time to consulting. Again, student satisfaction declines.

Summary and Conclusions

The purpose of this paper is to determine to what extent pedagogy affects student satisfaction as measured by student evaluations of teaching in Principles of Economics. The results suggest that pedagogy matters. Instructors do retain some control over their course evaluations. Increasing evaluation scores is possible through intensive use of homework, a more comprehensive final exam, and a higher percentage of class time devoted to instructor-student dialogue. However, use of the same text as the previous semester, increased efforts spent in consulting endeavors, intensive use of videos, and higher percentages of the final grade being determined by quizzes, homework, and midterm exams, will lead to lower ratings. Other choices, such as intensive use of a study guide, computer exercises, and non-text readings, prove to be inconsequential for student satisfaction, as does the percentage of the final grade being determined by written assignments. Students are indifferent as to the type of questions used in assessing their understanding. Professors who allocate more time to teaching, research, or administrative responsibilities do not seem to be placing themselves at a greater disadvantage.

While the findings of this study do suggest that pedagogy matters, they do not lend conclusive support to advocates of pedagogy changes, such as movements away from lecture and multiple-choice exams in introductory economics courses (Becker and Watts 2001). However, as mentioned previously with respect to the integration of computers into the classroom, some non-lecture forms of pedagogy (e.g., service learning, write-to-learn, experimental methods) either were in their infancy or not used at all at the time of TUCE III. Later research may reveal these forms of pedagogy to be effective means of improving SET scores.

Finally, we note that in addition to student satisfaction, pedagogical choices may influence cognitive outcomes, may be more or less well matched to student learning styles, or could have consequences for economic efficiency in their adoption. Clearly, effective teaching requires careful consideration of a variety of dimensions. It would be interesting to know to what extent pedagogy matters for these other aspects of classroom instruction.

References

American Association of Colleges and Schools of Business International. 1999. *Achieving Quality and continuous Improvement through Self-Evaluation and Peer Review: Standards for Accreditation* (Reprinted January 20). St. Louis.

Aigner, Dennis J., and Frederick D. Thum. 1986. "On Student Evaluation of Teaching Ability." *Journal of Economic Education* 17 (Fall): 243-265.

Anderson, Kathryn H., and John J. Siegfried. 1997. "Gender Differences in Rating the Teaching of Economics." *Eastern Economic Journal* 23 (Summer): 347-357.

Basow, S. A., and N. T. Silberg. 1987. "Student Evaluations of College Professors: Are Female and Male Professors Rated Differently?" *Journal of Educational Psychology* (79): 308-314.

Becker, William E. 1997. "Teaching Economics to Undergraduates." *Journal of Economic Literature* 35 (September): 1347-1373.

Becker, William E., and Walstad, Walter B. 1987. "Statistical Methods in Economic Education Research." In *Econometric Modeling in Economic Education Research,* edited by W. E. Becker, and W. B. Walstad. Boston, MA: Kluwer-Nijhoff.

Becker, William E., and Michael Watts. 2001. "Teaching Methods in U.S. Undergraduate Economics Courses." *Journal of Economic Education* (Summer): 269-279.

________. 1999. "How Departments of Economics Evaluate Teaching." *American Economic Review* 89 (May): 344-349.

Bejar, L. L., and K. O. Doyle. 1978. "Relationship of Curriculum Area and Course Format with Students' Ratings of Instruction." *American Educational Research Journal*15: 483-487.

Boex, Jameson, L. F. 2000. "Attributes of Effective Economics Instructors: An Analysis of Student Evaluations." *Journal of Economic Education* Summer 31: 211-227.

Bosshardt, William, and Michael Watts. 2001. "Comparing Student and Instructor Evaluations of Teaching." *Journal of Economic Education* 32 (Winter): 3-17.

Braskamp, Larry A., and John C. Ory. 1994. *Assessing Faculty Work: Enhancing Individual and Institutional Performance*. San Francisco: Jossey-Bass.

Feldman, K. A. 1993. „College Students' Views of Male and Female College Teachers: Part II—Evidence from Students' Evaluations of Their Classroom Teacher." *Research in Higher Education* 34: 151-211.

________. 1983. "Seniority and Experience of College Teachers as Related to Evaluations They Receive from Students." *Research in Higher Education* 18: 3-124.

________. 1976. "Grades and College Students' Evaluations of Their Courses and Teacher." *Research in Higher Education* 4: 69-111.

________. "Research Productivity and Scholarly Accomplishments of College Teachers as Related to Their Institutional Effectiveness: A Review and Exploration," Research in Higher Education, 1987, 26, 227-291.

Gomez-Mejia, L., and Balkin, D. 1992. "The Determinants of Faculty Pay: An Agency Theory Perspective." *Academy of Management Journal* 35: 921-955

Gramlich, Edward M., and Glen A. Greenlee. 1993. "Measuring Teaching Performance." *Journal of Economic Education* 24 (Winter): 3-13.

Greenwald, A. G., and G. M. Gillmore. 1997. "No Pain, No Gain? The Importance of Measuring Course Workload in Student Ratings of Instruction." *Journal of Educational Psychology* 89: 743-751.

Hanushek, E. A. 1979. "Conceptual and Empirical Issues in the Estimation of Educational Production Functions." *Journal of Human Resources* 14: 351-388.

Herzberg, Frederick. 1987. "One More Time: How Do You Motivate Employees?" *Harvard Business Review* (September/October) 65: 109-120.

Johnston, C., I. McDonald, and R. Williams. 2001. "The Scholarship of Teaching Economics." *Journal of Economic Education* 32 (Summer): 195-201.

Kelley, A. C. 1975. "The Student as a Utility Maximizer." *Journal of Economic Education* 6 (Spring): 82-92.

Leeds, Michael, William Stull, and Jilleen Westbrook. 1998. "Do Changes in Classroom Techniques Matter? Teaching Strategies and Their Effects on Teaching Evaluations." *Journal of Education for Business* 74 (November/December): 75-78.

Lumsden, K. G. 1974. "The Information Content of Student Evaluation of Faculty and Courses." In *Efficiency in Universities: The LaPaz Papers*, edited by K. Lumsden. Amsterdam and New York: Elsevier Scientific: 175-204.

Lumsden, K. G., and Scott, A. 1983. "The Efficacy of Innovative Teaching Techniques in Economics: The U.K. Experience." *American Economic Review* (May): 13-17.

Manahan, J. 1983. "An Educational Production Function for Principles of Economics." *Journal of Economic Education* 14 (Spring): 11-16.

Mason, P. M., J. W. Steagall, and M. M. Fabritus. 1995. "Student Evaluations of Faculty: A New Procedure for Aggregate Measures of Performance." *Economics of Education Review* 14: 403-416.

McKenzie, R. B., and R. J. Staff. 1974. *An Economic Theory of Learning: Student Sovereignty and Academic Freedom*. Blacksburg, VA: University Publications.

Mehdizadeh, M. 1990. "Loglinear Models and Student Course Evaluations." *Journal of Economic Education* 21 (Winter): 7-21.

Morgan, W. D., and J. D. Vasche. 1978. "An Educational Production Function Approach to Teaching Effectiveness and Evaluation." *Journal of Economic Education* 9 (Spring): 123-126.

Needham, Douglas. 1978. "Student Effort, Learning and Course Evaluation." *Journal of Economic Education* 9 (Fall): 35-43.

Nelson, J. P., and K. A. Lynch. 1984. "Grade Inflation, Real Income, Simultaneity, and Teaching Evaluations." *Journal of Economic Education* 15 (Winter): 21-37.

Saunders, Phillip. 1994. *The TUCE III Data Set: Background Information and File Codes*. New York: National Council on Economic Education.

Siegfried, John J., Phillip Saunders, Ethan Stinar, and Hao Zhang. 1996. "Teaching Tools: How Is Introductory Economics Taught in America?" *Economic Inquiry* 34 (January): 182-186.

Southern Association of Colleges and Schools. 1996. *Criteria for Accreditation*. Commission on Colleges.

Stechlein, J. F. 1960. *Encyclopedia of Educational Research*, edited by C. Harris. New York: MacMillian Publishing Company: 286-287.

Stratton, Richard W., Steven C. Myers, and Randall H. King. 1994. "Faculty Behavior, Grades, and Student Evaluations." *Journal of Economic Education* 25 (Winter): 5-15.

Tang, Thomas Li-Ping. "Teaching Evaluation at a Public Institution of Higher Education: Factors Related to the Overall Teaching Effectiveness." *Public Personnel Management* 25: 379-384.

Watts, M. and G. Lynch. 1989. "The Principles Course Revisited." *American Economic Review* 79 (May): 236-241.

White, Lawrence J. "Efforts by Departments of Economics to Assess Teaching Effectiveness." *Journal of Economic Education* (Winter) 26: 81-85.

Wilson, Robin. 1998. "New Research Casts Doubt on Value of Student Evaluations of Professors." *Chronicle of Higher Education* 44 (January 16): A12.

Yunker, J. A., and Marlin, J. W. 1984. "Performance Evaluation of College and University Faculty: An Economic Perspective." *Education Administration Quarterly* 20: 9-37.

Zietz, J., and Howard H. Cochran, Jr. 1998. "How Much Does Teaching Methodology Matter for Learning Principles of Economics?" *The Kentucky Journal of Economics and Business* 17: 39-54.