

# **Student Course Evaluations: Key Determinants of Teaching Effectiveness Ratings in Accounting, Business and Economics Courses at a Small Private Liberal Arts College**

**Timothy M. Diette, Washington and Lee University  
George W. Kester, Washington and Lee University**

## **ABSTRACT**

*In this study, we examined the individual-level student evaluations of courses and teachers (SET) used in the Williams School of Commerce, Economics and Politics at Washington and Lee University to assess the most important determinants of teaching effectiveness ratings by students. Using factor analysis, we found that the primary driver affecting student ratings of teaching effectiveness is a factor related to the quality of instruction. We also found that a factor associated with course workload is not related to teaching effectiveness ratings, which goes against the belief that it is difficult to receive high teaching effectiveness ratings in courses with heavy workloads. Using multiple variable regression, we found that the most important determinants of teaching effectiveness ratings are clear communication of the main points of lectures/seminars, meaningful and conscientious evaluation of student work, and professor enthusiasm. Using the same methods, we also examined the determinants of whether students would recommend professors to other students.*

## **INTRODUCTION**

An important input into faculty merit, tenure, and promotion reviews is the result of student evaluations of courses and teachers (SET hereafter). Although there other methods of evaluating teaching effectiveness, such as peer visitation, student interviews, and alumni surveys, SETs are probably the most important and certainly the most commonly-used basis for evaluating teaching effectiveness.

In the Williams School of Commerce, Economics and Politics at Washington and Lee University, an SET is administered at the end of each course. Although in some courses, students complete the questionnaires electronically (via the internet using Sakai, SurveyMonkey or Blackboard), the majority of instructors prefer to administer the surveys using “old school” paper questionnaires. In both cases, the questionnaires are completed and submitted anonymously. The completed questionnaires are tabulated and the results are provided to the faculty member and his or her department head. The tabulations are kept on file and are required as part of faculty annual performance reviews and tenure and promotion files.

As a matter of policy, the SET contains 17 common questions, 14 of which are based upon standard Likert-type scales from 1 to 5. These questions (hereafter identified by “Q”) assess student perceptions of various aspects of teaching such as instructor preparation,

enthusiasm, approachability, accessibility, course organization and workload. Two of the questions are open-ended for students' written comments regarding the course and professor.<sup>1</sup> The SET common questions are shown in the Appendix.

The purpose of our research project is to determine which aspects of the SETs are most related to overall teaching effectiveness as perceived by students in two questions in our SET: Q13 which asks students to "rate the professor's overall teaching effectiveness" and Q14 which asks if students "would recommend this professor to other students?"

Our study is based upon students and SET's in a somewhat unique educational setting. Washington and Lee University (W&L) is a small highly selective private institution and the only top-tiered national liberal arts college with an AACSB-accredited school of business.<sup>2</sup> Founded in 1749, W&L has a long history of providing business education at the undergraduate level. The Williams School of Commerce, Economics and Politics (hereafter the Williams School) was accredited by the AACSB in 1927. Classes at W&L are small, generally ranging in size from 15 to 20 students. High quality undergraduate teaching is the highest priority of our faculty, with seminar-style classes emphasizing classroom discussions and presentations being preferred to lecture-oriented classes. Therefore, the expectations of our students regarding the attributes of an effective teacher may be quite different than those found at larger universities.

## **RELATED LITERATURE**

The academic literature on SETs is quite extensive, ranging from examining the impact on overall teaching ratings of grading leniency (Greenwald and Gilmore, 1997, McPherson, 2006, Moore, 2006, and Constand and Pace, 2012), differences across academic disciplines (Feldman, 1978, Cashin, 1990, Obenchain, Abernathy and West, 2001, Felton, Mitchell and Stinson, 2004a, and Constand and Pace, 2012) and even instructor physical attractiveness (Felton, Mitchell and Stinson, 2004b, Hamermesh and Parker, 2005, Riniolo et al., 2006, Feng & Weber, 2009, and Constand and Pace, 2012).

Using exploratory factor analysis, Barth (2008) examined SET responses from the College of Business Administration of Georgia Southern University to identify five factors (he termed quality of instruction, course rigor, grades, level of student interest, and instructor helpfulness) that he related to the SET questionnaire overall teaching effectiveness rating using multiple regression analysis. He found that the most important determinant of the overall teaching rating was the quality of instruction followed by instructor helpfulness. He found a small but negative relationship between course rigor and the overall rating. Our study also utilizes both factor analysis and multiple variable regression with the individual assessment questions to identify the key determinants of teaching effectiveness as perceived by our students. Whereas Barth's level of observation was at the course section level (167 classes), we are able to observe individual student evaluations (860 SET's).

## **DATA AND METHOD**

We are interested in understanding the characteristics of professors that are associated with higher teacher effectiveness ratings by our students as well as a higher likelihood of recommending the professor to other students. For our study, we collected the individual level student course evaluations from 19 professors across three disciplines in the Williams School: accounting, business administration, and economics for courses taken during the Fall 2012

semester. We also included sections of our interdepartmental statistics course required of all Williams School majors. The course evaluations (shown in the Appendix) have the same first 17 questions for all professors. Some professors add additional questions for their students, but these are not included in our analysis. Our analysis sample contains 860 individual course evaluation questionnaires (observations).<sup>3</sup>

Table 1 shows the questions in the Williams School SET, with the overall results from the 860 questionnaires for each question. The first column of data in Table 1 reports the mean response on a 1 to 5 scale (except for Q15 where students report the mean number of hours spent outside of class on the course) for all students in the sample to each of the questions used in the analysis. The second column reports the means for students in the core (required) courses while the last column reports the mean response for students in elective courses. The table is organized into three panels. The first panel contains the two dependent (outcome) variables that will be explored: Q13 as our measure of overall teaching effectiveness and Q14 is the measure of whether a student would recommend the professor to other students. The overall evaluations are quite high with the mean overall teaching effectiveness of 4.30 across all observations and 4.43 on recommending the professor.<sup>4</sup> As might be expected, both measures have higher mean scores for elective courses than for the core courses.

The first three independent variables (Q1, Q2 and Q3) are intended to examine student perceptions of the design and organization of the course. The remaining questions focus on the students' perceptions of their professor during class, in personal interactions, and their treatment of the students' assignments.<sup>5</sup> All questions have means greater than 4.00 with the exception of Q3. Q3 asks the students about the relative workload on a 1 to 5 scale. With a mean score of 3.33 for all courses, students perceive the workload to be higher than other courses they are taking or have taken. The mean score is slightly higher for elective courses.

We analyze the data using two alternative approaches. We first use exploratory factor analysis and subsequent regressions with the factors to analyze our data. This is the same methodology employed by Barth (2008) in his analysis of student evaluations from courses at Georgia Southern University. Due to potential concerns regarding the explanatory variables (Q1-Q3, Q5-Q12, and Q15 in our SET) being highly correlated, factor analysis reduces the dimensions of the data. The reduced number of dimensions can also make the results easier to interpret as the analysis groups together questions that conceptually measure similar aspects of teaching. The results from the first stage suggest the appropriate number of factors to be retained for the multiple regression and enables analysis of the factor loadings or correlation coefficients between the variables and factors. If several variables are highly correlated, then the subsequent regression suffers from multicollinearity and the parameter estimates for these variables may not be valid. However, a factor that represents several highly correlated questions provides a clearer interpretation about the relative association of that factor with teaching effectiveness. After generating each factor, the multiple variable regression includes the factors and additional controls outside of the SET.<sup>6</sup>

In the second approach, we use multiple ordinary least squares regressions with the original questions as independent variables with the additional control variables. We then test for whether multicollinearity is a problem with the instrument used in our analysis. Based on variance inflation factor (VIF) scores, we determine that this specification does not suffer from multicollinearity. Therefore, while factor analysis is still useful for one level of interpretation, these results may provide additional insights into characteristics associated with effective professors as determined by students' perceptions.

**Table 1: Student Evaluation Instrument and Mean Scores**

<b>Question</b>		<b>All Courses (Mean)</b>	<b>Core Courses (Mean)</b>	<b>Elective Courses (Mean)</b>
<b>Dependent Variables</b>				
Q13	Rate the professor's overall teaching effectiveness.	4.30	4.22	4.48
Q14	Would you recommend this professor to other students?	4.43	4.34	4.60
<b>Independent Variables</b>				
Q1	Was the course well organized?	4.58	4.53	4.67
Q2	Were student responsibilities/requirements made clear at the beginning of the term?	4.63	4.57	4.74
Q3	How did the workload in this course compare to other courses you have taken or are taking?	3.33	3.26	3.48
Q5	The main points of lectures/seminars were clearly communicated.	4.39	4.36	4.46
Q6	Was the professor well prepared for course meetings?	4.76	4.75	4.77
Q7	Was the professor enthusiastic about teaching the course?	4.68	4.60	4.85
Q8	How accessible was the professor for student meetings about the course?	4.55	4.50	4.68
Q9	Was the professor approachable?	4.67	4.62	4.78
Q10	Did the professor return assignments in a timely manner?	4.37	4.42	4.28
Q11	The professor evaluated my work in a meaningful, conscientious manner.	4.39	4.33	4.54
Q12	Did the professor treat you with respect?	4.82	4.79	4.86
Q15	On average, how many hours per week did you spend outside of class for this course?	5.20	4.95	5.74
<b>Additional Course Control Variables</b>				
Elective Course		32%	0%	100%
Business Course		30%	27%	39%
Accounting Course		18%	20%	16%
Observations		860	584	276
Notes: The reference group (omitted category) for Elective Course is Required Course. The reference group for the Business Course and Accounting Course variables is Economics Courses and the Interdepartmental Statistics Course.				

## RESULTS

### Factor Analysis Results

The first decision required within factor analysis is to determine the appropriate number of factors to include. Researchers typically examine the eigenvalues of factors and retain the number of factors with eigenvalues greater than one. This method is commonly referred to as the Kaiser rule. A second common approach is to examine the scree plot, a plot of the eigenvalues. Under this method, the number of factors is based on when the curve of eigenvalues flattens. These criteria suggest that the appropriate number is either two or three factors. The results for the remainder of the analysis are consistent whether two or three factors are retained. For ease of interpretation, we report the results for three factors.

Table 2 reports the rotated factor pattern and factor loadings or correlations with the original set of variables. All factor loadings of 0.40 or higher are in bold and italics. The three factors divide quite neatly across three general characteristics. The first factor is highly correlated with measures associated with course design, management, and the enthusiasm within the classroom. We label this factor *quality of instruction* (following Barth's nomenclature). The second factor is associated with the professor being deemed approachable, accessible for meetings, and having treated students with respect. We call this second factor *approachability*. The third and final factor which we call *workload* captures the two questions related to the demands of the course: the relative workload and the average number of hours per week outside of class.

**TABLE 2: Rotated Factor Pattern and Factor Loadings**

	Factor 1 ( <i>Quality of Instruction</i> )	Factor 2 ( <i>Approachability</i> )	Factor 3 ( <i>Workload</i> )
Q1 (Course Organized)	<b><i>0.78</i></b>	-0.05	0.01
Q2 (Clear Responsibilities)	<b><i>0.49</i></b>	0.24	-0.05
Q3 (Workload)	0.02	-0.03	<b><i>0.93</i></b>
Q5 (Clear Main Points)	<b><i>0.81</i></b>	-0.12	-0.03
Q6 (Prepared for Class)	<b><i>0.80</i></b>	-0.07	0.02
Q7 (Enthusiastic)	<b><i>0.40</i></b>	0.17	0.15
Q8 (Accessible for Meetings)	0.07	<b><i>0.61</i></b>	0.07
Q9 (Professor Approachable)	-0.17	<b><i>0.99</i></b>	-0.03
Q10 (Return Work Promptly)	0.38	0.24	0.00
Q11 (Evaluation of Work)	0.28	0.39	0.01
Q12 (Treated with Respect)	0.21	<b><i>0.43</i></b>	-0.12
Q15 (Average Hours Per Week)	-0.07	0.03	<b><i>0.51</i></b>

*Note.* Values in bold and italics are the factor loadings greater than 0.40.

### Regression Analysis Using Factor Variables

The factors are then entered as independent variables in a multiple regression along with

a dummy or indicator variable for whether the course is an elective class (with core courses as the reference group) and indicator variables for whether the course is in Business Administration or Accounting (with courses in Economics and statistics as the reference group). The results for perceived teaching effectiveness (Q13) as the dependent variable are reported in Table 3. The results for how likely the student is to recommend the professor to other students (Q14) are shown in Table 4. Both tables contain three columns of parameter estimates. The first column of estimates is for the entire sample of SET's. The second and third columns correspond to results when the data are stratified by core versus elective status which allow us to examine whether the relationships between the factors and the dependent variables differs by the type of course. The results for core courses are shown in the second column of results and the results for the estimates for the elective courses are in the third column.

**TABLE 3: Teaching Effectiveness (Q13) Regression Results Using Factor Variables**

	(1) All Courses		(2) Core Courses		(3) Elective Courses	
	Parameter Estimate	<i>t</i>	Parameter Estimate	<i>t</i>	Parameter Estimate	<i>t</i>
Factor1 (Quality of Instruction)	0.604	12.56***	0.604	10.17***	0.604	5.69***
Factor2 (Approachability)	-0.011	-0.23	-0.000	-0.00	-0.063	-0.56
Factor3 (Workload)	0.021	0.60	0.024	0.54	0.011	0.27
Elective Course	0.145	2.17**				
Business Course	0.006	0.06	-0.054	-0.38	0.120	1.36
Accounting Course	0.102	2.37**	0.086	2.16**	0.146	1.56
Intercept	4.236	111.61***	4.257	105.11***	4.341	46.60***
<i>N</i>	860		584		276	

*Notes:* \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ; The dependent variable was response to Question 13 “Rate the professor’s overall teaching effectiveness.”

**TABLE 4: Recommend the Professor (Q14) Regression Results Using Factor Variables**

	(1) All Courses		(2) Core Courses		(3) Elective Courses	
	Parameter Estimate	<i>t</i>	Parameter Estimate	<i>t</i>	Parameter Estimate	<i>t</i>
Factor1 (Quality of Instruction)	0.506	9.08***	0.584	9.36***	0.302	3.50***
Factor2 (Approachability)	0.129	2.08*	0.066	1.19	0.295	2.28**
Factor3 (Workload)	-0.010	-0.22	-0.016	-0.33	-0.008	-0.14
Elective Course	0.130	1.52				
Business Course	0.002	0.02	-0.137	-0.85	0.277	2.75**
Accounting Course	0.110	1.53	0.041	0.70	0.303	2.78**
Intercept	4.363	69.89***	4.413	77.78***	4.354	60.05***
<i>N</i>	860		584		276	

*Notes:* \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ; The dependent variable was response to Question 14 “Would you recommend this professor to other students?”

In Table 3, all three sets of results find a strong positive association between the *quality of instruction* and teaching effectiveness ratings. Interestingly, neither the factor variables related to *approachability* nor *workload* appears to be related to perceived teaching effectiveness. One possible explanation for the insignificance of *approachability* is that most of the professors in the small liberal arts setting of the Williams School have “open door” office policies and may all be viewed as quite approachable. Among the other control variables, elective courses are positively related with teaching effectiveness. As previously mentioned, professors in elective courses receive higher evaluations of teaching effectiveness. In these “self-selected” courses, students are able to choose both the course and the professor.

The results for Table 4 are broadly consistent with the results in Table 3. However, there are several notable differences. While the *approachability* factor was statistically unrelated with teaching effectiveness, *approachability* is positively related to whether a student will recommend a particular professor to other students. This relationship held for all courses and all electives, but not for core courses. Professors teaching elective courses are no more likely to be recommended than those in core courses (although the *p-value* is 0.145). And among elective courses, students in business administration and accounting courses were more likely to recommend their professors to other students than those students in economics or statistics courses after controlling for the three factors.

### **Regression Analysis Using Original Variables**

There is a two-fold justification for carrying out the factor analysis performed above: interpretability of the results and as a way of addressing the potential problem of multicollinearity. Next, we estimated multiple regressions including all of the original questions used in the first stage of the factor analysis directly in the regressions. We tested for the presence of multicollinearity by calculating the variance inflation factor (VIF) scores. The VIF scores range from 1.58 to 1.61, suggesting that multicollinearity may not be a significant concern with our sample.

The results from the multiple regressions using the original variables are qualitatively similar to the multiple regression analysis that utilizes the factor scores and are reported in Tables 5 and 6 for teaching effectiveness and likelihood of recommending the professor, respectively.

In Table 5, we can analyze the relative importance of the components (Q1, Q2, Q5, Q6 and Q7) of the *quality of instruction* factor. Course organization (Q1), providing clear main points of the lectures (Q5), and enthusiasm (Q7) are all statistically significantly associated with higher ratings of teacher effectiveness. Making the responsibilities clear at the start of the term (Q2) and being prepared for class (Q6) were not directly related. As acknowledged by the proponents of factor analysis, this is an example where the results are not easily put into practice. It is not entirely clear how a professor would be unprepared for class but still clearly emphasizes the main points and has a well-organized course. As previously discussed, the *approachability* factor was statistically insignificant in the factor analysis regression for teaching effectiveness. However, its components, accessibility for meetings (Q8) and being treated with respect (Q12) are both insignificant here and being approachable (Q9) is positively related to higher teaching effectiveness ratings in the full sample and core courses, but not elective courses. One advantage provided by this method is shown by Q11, the parameter estimate for whether the professor evaluates the student work in a meaningful and conscientious manner. In our factor analysis, it

was only weakly associated with the *approachability* factor, a factor that was statistically insignificant. In our multiple regression, we find that Q11 is positive and significantly related to higher teaching effectiveness ratings.

Finally, a higher relative workload (Q3) in elective courses was associated with lower teaching effectiveness, but average hours per week (Q15) was positively associated with teaching effectiveness in the full sample and both subsamples.

The results in Table 6, those for whether a student would recommend a professor to others, are quite similar to those for teaching effectiveness. One key result is that professors in elective courses with larger relative workloads (Q3) are less likely to be recommended. In addition, this is no longer offset by positive effect from the average hours of work per week (Q15) for either core or elective courses. Consistent with the findings in Table 5, students' perceptions of the meaningful and conscientious evaluation of their work (Q11) is positively related with whether they would recommend a professor to their peers.

**TABLE 5: Teaching Effectiveness (Q13) Regression Results Using Original Variables**

	(1) All Courses		(2) Core Courses		(3) Elective Courses	
	Parameter Estimate	<i>t</i>	Parameter Estimate	<i>t</i>	Parameter Estimate	<i>t</i>
Q1 (Course Organized)	0.179	2.89**	0.204	3.12***	0.126	1.06
Q2 (Clear Responsibilities)	0.067	1.33	0.058	0.98	0.064	0.83
Q3 (Workload)	-0.013	-0.40	-0.006	-0.15	-0.039	-1.84*
Q5 (Clear Main Points)	0.400	9.52***	0.428	7.77***	0.306	7.91***
Q6 (Prepared for Class)	0.017	0.35	-0.050	-0.82	0.149	1.76
Q7 (Enthusiastic)	0.254	4.11***	0.270	3.86***	0.173	4.07***
Q8 (Accessible for Meetings)	-0.049	-1.13	-0.037	-0.74	-0.089	-1.54
Q9 (Professor Approachable)	0.088	1.84*	0.126	1.98*	-0.041	-0.39
Q10 (Return Work Promptly)	-0.032	-1.25	-0.051	-1.15	-0.002	-0.05
Q11 (Evaluation of Work)	0.216	4.99***	0.165	3.40***	0.328	4.52***
Q12 (Treated with Respect)	-0.034	-0.47	-0.033	-0.39	0.033	0.27
Q15 (Avg. Hours Per Week)	0.017	3.32***	0.016	1.91*	0.021	2.66**
Elective Course	0.067	1.25				
Business Course	-0.018	-0.23	-0.067	-0.58	0.086	1.46
Accounting Course	0.073	1.88*	0.097	2.22**	0.052	0.88
Intercept	-0.760	-1.66	-0.653	-1.15	-0.369	-1.00
<i>N</i>	860		584		276	

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ; The dependent variable was response to Question 13 "Rate the professor's overall teaching effectiveness."

**TABLE 6: Recommend the Professor (Q14) Regression Results Using Original Variables**

	(1) All Courses		(2) Core Courses		(3) Elective Courses	
	Parameter Estimate	<i>t</i>	Parameter Estimate	<i>t</i>	Parameter Estimate	<i>t</i>
Q1 (Course Organized)	0.120	1.96*	0.116	1.99*	0.112	0.97
Q2 (Clear Responsibilities)	0.065	1.10	0.088	1.35	-0.080	-1.37
Q3 (Workload)	-0.030	-0.75	-0.013	-0.28	-0.091	-2.96**
Q5 (Clear Main Points)	0.326	7.54***	0.382	6.99***	0.146	2.84**
Q6 (Prepared for Class)	0.062	0.85	0.133	1.40	-0.027	-0.24
Q7 (Enthusiastic)	0.193	3.30***	0.150	2.28**	0.357	4.16***
Q8 (Accessible for Meetings)	-0.019	-0.44	-0.046	-0.85	0.092	1.37
Q9 (Professor Approachable)	0.159	3.49***	0.167	2.91**	0.103	0.87
Q10 (Return Work Promptly)	0.002	0.05	-0.025	-0.55	0.041	0.96
Q11 (Evaluation of Work)	0.208	3.96***	0.181	3.44***	0.296	2.92**
Q12 (Treated with Respect)	0.101	1.02	0.068	0.61	0.171	1.18
Q15 (Avg. Hours Per Week)	0.013	1.61	0.004	0.39	0.016	1.29
Elective Course	0.078	1.10				
Business Course	-0.006	-0.05	-0.135	-0.89	0.209	2.45**
Accounting Course	0.092	1.33	0.056	0.92	0.189	2.26**
Intercept	-1.128	-1.90**	-1.093	-1.60*	-0.953	-1.55
<i>N</i>	860		584		276	

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ; The dependent variable was response to Question 14 “Would you recommend this professor to other students?”

## DISCUSSION

SET questionnaires are completed by students at one particular time during or after a course and reflect *student perceptions* of teaching effectiveness and the other dimensions of teaching, not necessarily *actual* teaching effectiveness. As discussed by Moore (2006), some faculty members are uncomfortable with SETs, believing that they may reflect popularity rather than actual teaching effectiveness. Others are concerned that teaching effectiveness ratings by students are influenced by course difficulty and grading, a concern supported by the research findings of Greenwald and Gillmore (1997), Barth (2008) and Constand and Pace (2013). As discussed by Stapleton and Murkison (2001), some faculty members are concerned that differences in course subject matter can affect student perceptions and teaching evaluation ratings. This concern is also consistent with the recent findings of Constand and Pace (2013), who found from their examination of teaching ratings contained in RateMyProfessor.com that finance professors and their courses received significantly lower overall teaching ratings than other business disciplines. They also found a positive relationship between the attractiveness of professors and teaching ratings.<sup>7</sup>

Notwithstanding these and other limitations of SET's, several conclusions can be drawn from our study. According to our factor analysis, the primary determinant of student ratings of teaching effectiveness is the *quality of instruction*. This result is similar to the findings of Barth (2008). Both *quality of instruction* and the factor related to *approachability* (except in the case

of core courses) are the primary determinants of whether the professor is recommended to other students. Interestingly, the factor related to *workload* was not related to either measure of teaching quality. This finding, which is inconsistent with Barth (2008) who found a negative relationship between the factor he labeled “course rigor” and teaching quality, goes against the belief that it is difficult to receive high teaching ratings in courses with heavy workloads.

Our results using multiple variable regressions of the individual SET questions indicate that the most important determinants of teaching effectiveness ratings are clear communication of the main points of lectures/seminars, evaluation of student work in a meaningful and conscientious manner, and instructor enthusiasm. Other determinants are course organization and approachability. Workload compared to other courses is not related to teaching effectiveness ratings in our entire sample or in core courses. However, in the case of elective courses, workload compared to other courses is negatively related to both teaching effectiveness ratings and whether the professor is recommended to other students. This result might suggest that students prefer elective courses with lower than average workloads and therefore it is more difficult to receive high teaching effectiveness ratings in electives with heavier workloads.

We had mixed results regarding the average number of hours spent outside class for courses. The hours spent outside class is positively related to teaching effectiveness. One interpretation of this finding is that students spend more time on courses that are effectively taught. Using the second dependent variable, whether a professor is recommended to other students, the average number of hours spent outside class is unrelated to whether the professor is recommended.

A question pertaining to perceived course difficulty is not included among the questions in the Williams School’s SET. Course difficulty was included along with questions related to workload and student effort in Barth’s “course rigor” factor.<sup>8</sup> Although courses with heavy workloads that require lots of hours outside class can also be difficult, it is also possible for courses to have heavy workloads but not be especially difficult. Therefore, a question related to perceived course difficulty would be a useful addition.

The SET used in the Williams School at W&L also does not include questions pertaining to the sex and class year (sophomore, junior, etc.) of the student, nor does it solicit information regarding the student’s grade point average (GPA) or expected grade in the course being evaluated. As we conducted this research project we realized that a logical extension of our research would be to revise the Williams School SET to solicit this information so that we can explore whether the determinants of teaching effectiveness ratings at the Williams School vary accord to sex, class year and grade.

Our results, of course, are specific to the William School’s particular SET and perceived determinants of overall teaching effectiveness by our students. As previously mentioned, W&L is a unique educational institution -- a small highly selective private liberal arts college with an accredited school of business. If the Williams School SET were administered in a university with a different teaching culture, the results may be different than found in our study.

## ENDNOTES

<sup>1</sup> “Professor” refers to all professors and course instructors regardless of their rank.

<sup>2</sup> In 2013, for example, Washington and Lee University ranked number 14 among National Liberal Arts Colleges in *U.S. News & World Report’s* Best College Rankings.

<sup>3</sup> Out of 964 SETs, 104 were missing between one and two responses. These evaluations are dropped from the sample, leaving 860 individual observations.

<sup>4</sup> The correlation between Q13, which asks students to “rate the professor’s overall teaching effectiveness” and Q14 which asks if students “would recommend this professor to other students” is 0.78.

<sup>5</sup> We do not make use of question 4 of the SET which asks whether students would recommend the course without specific reference to the professor.

<sup>6</sup> A more detailed discussion of factor analysis is provided by Barth (2008).

<sup>7</sup> See Germain and Scandura (2005) for an excellent discussion of the validity of SETs as instruments for measuring teaching effectiveness.

<sup>8</sup> Barth’s (2008) factor labeled “course rigor” included the results of SET questions related to course difficulty and workload, as well as the amount of effort put into the course, how often students sought outside help with the course, and the degree to which they were challenged in the course.

## REFERENCES

Barth, M. M. “Deciphering Student Evaluations of Teaching: A Factor Analysis Approach,” *Journal of Education for Business*, 84 (2008), pp. 40-46.

Cashin, W.E. “Students Do Rate Different Academic Fields Differently,” in *Student Ratings of Instruction: Issues for Improving Practice*, M. Theall and J. Franklin (editors), Jossey-Bass, San Francisco, 1990, pp. 113-122.

Constand, R. L., and R. D. Pace. “Student Evaluations of Finance Faculty: Perceived Difficulty Means Lower Faculty Evaluations,” paper presented at the Academy of Business Education/Financial Education Association 2012 Conference.

Feldman, K. A. “Course Characteristics and College Student’s Ratings of Their Teachers: What We Know and What We Don’t,” *Research in Higher Education*, 21 (1978), pp. 199-242.

Felton, J., J. Mitchell, and M. Stinson. “Web-Based Student Evaluations of Professors: The Relationships Between Perceived Quality, Easiness and Sexiness,” *Assessment and Evaluation in Higher Education*, 29 (2004), pp. 91-108.

Felton, J., J. Mitchell, and M. Stinson. “Cultural Differences in Student Evaluations of Professors,” Proceedings of the Academy of Business Education/Financial Education Association 2004 Conference, *Journal of the Academy of Business Education*, 6 (retrieved from <http://abeweb.org/proceedings/proceedings04/felton2.pdf>).

Feng, S., and D. Webber. “Turning up the Heat on Online Teaching Evaluations: Does

‘Hotness’ Matter?” *Teaching of Psychology*, 36 (2009), pp. 189-193.

Germain, M. L., and T. Scandura. “Grade Inflation and Student Individual Differences as Systematic Bias in Faculty Evaluations,” *Journal of Instructional Psychology*, 32 (2005), pp. 58-67.

Greenwald, A. G., and G. M. Gilmore. “Grading Leniency is a Removable Contaminant of Student Ratings,” *American Psychologist*, 52 (1997), pp. 1209-1219.

Hamermesh, D. S., and A. M. Parker. “Beauty in the Classroom: Instructor’ Pulchritude and Putative Pedagogical Productivity,” *Economics of Education Review*, 24 (2005), pp. 369-376.

McPherson, M. “Determinants of How Students Evaluate Teachers,” *Journal of Economic Education*, 37 (2006), pp. 3-20.

Moore, T. “Teacher Evaluations and Grades: Further Evidence,” *Journal of American Academy of Business*, 9 (2006), pp. 58-62.

Obenchain, K. M., T. V. Abernathy and L. R. Wiest. “The Reliability of Students’ Ratings of Faculty Teaching Effectiveness,” *College Teaching*, 49 (2001), pp. 100-104.

Riniolo, T. C., K. C. Johnson, T. R. Sherman, and J. A. Misso. “Hot or Not: Do Professors Perceived as Physically Attractive Receive Higher Student Evaluations?” *The Journal of General Psychology*, 133 (2006), pp. 19-35.

Stapleton, R., and Murkinson, G. “Optimizing the Fairness of Student Evaluations: A Study of Correlations Between Instructor Excellence, Study production, Learning Production, and expected Grades,” *Journal of Management Education*, 25 (2001), pp. 269-291.

*U.S. News & World Report’s Best College Rankings, National Liberal Arts Colleges* (retrieved from <http://colleges.usnews.rankingsandreviews.com/best-colleges/rankings/national-liberal-arts-colleges?int=a557e6>).

**APPENDIX: Student Evaluation of Teaching Form used at the Williams School of Commerce Economics and Politics, Washington and Lee University**

**COURSE EVALUATION**

*Please do not write your name or any other identification on this form*

Term \_\_\_\_\_ Year \_\_\_\_\_  
 Course \_\_\_\_\_ Section \_\_\_\_\_  
 Professor \_\_\_\_\_

Please score the following questions (you may comment on them below, if you desire):

1. Was the course well organized?  
 (Almost never)      1      2      3      4      5      (Almost always)      No opinion
2. Were student responsibilities/requirements made clear at the beginning of the term?  
 (No, not at all)      1      2      3      4      5      (Yes, very clear)      No opinion
3. How did the workload in this course compare to other courses you have taken or are taking?  
 (Much less)      1      2      3      4      5      (Much more)      No opinion
4. Would you recommend this **course** to other students?  
 (Not recommended)      1      2      3      4      5      (Highly recommended)      No opinion
5. The main points of lectures/seminars were clearly communicated.  
 (Almost never)      1      2      3      4      5      (Almost always)      No opinion
6. Was the professor well prepared for course meetings?  
 (Almost never)      1      2      3      4      5      (Almost always)      No opinion
7. Was the professor enthusiastic about teaching the course?  
 (Almost never)      1      2      3      4      5      (Almost always)      No opinion
8. How accessible was the professor for student meetings about the course?  
 (Never available)      1      2      3      4      5      (Available regularly)      No opinion
9. Was the professor approachable?  
 (No, seldom)      1      2      3      4      5      (Yes, always)      No opinion
10. Did the professor return assignments in a timely manner?  
 (Almost never)      1      2      3      4      5      (Almost always)      No opinion
11. The professor evaluated my work in a meaningful, conscientious manner.  
 (Strongly disagree)      1      2      3      4      5      (Strongly agree)      No opinion
12. Did the professor treat you with respect?  
 (No, seldom)      1      2      3      4      5      (Yes, always)      No opinion
13. Rate the professor's overall teaching effectiveness.  
 (Exceptionally low)      1      2      3      4      5      (Exceptionally high)      No opinion
14. Would you recommend this **professor** to other students?  
 (Not recommended)      1      2      3      4      5      (Highly recommended)      No opinion
15. On average, how many hours per week did you spend outside of class for this course? \_\_\_\_\_
16. Please comment on the course.
17. Please comment on the professor.