The Impact of Relative Grade Expectations on Student Evaluation of Teaching

Clifford Nowell

Abstract

It is commonly accepted that student evaluation of teaching (SET) ratings are influenced by expected grades, and that faculty are able to 'buy' higher SET ratings by giving higher grades. Researchers have questioned whether there are limits to the ability to buy grades due to the possibility that students reward teachers for their relative grade as opposed to their absolute grade.

In this paper we use SET data to investigate the relationship between SET ratings and relative grades. Similar to the prior literature, we find an indirect relationship between SET scores and historical grade performance averages (GPAs) but, we find the opposite result to be true when we examine the relationship between SET scores and expected grades earned by peers. Contrary to recent literature that suggests limits exist to an instructor's ability to purchase high SET scores when relative grades are considered, we find that the incentives to lower grading standards and buy higher SET ratings may actually be greater than has been thought in the past.

Introduction

Economists and sociologists have long recognised that satisfaction may depend on one's own circumstances or one's circumstances relative to a reference group (Becker, 1974). Indeed, a host of economists have explored the notion that satisfaction with one's own level of consumption and saving is often referenced by that of one's neighbours (Veblen, 1899; Duesenberry, 1949). More recently, Luttmer (2005) presented a set of convincing data that self-reported satisfaction may come from one's relative income position rather than one's absolute income position. He found that the inverse relationship between satisfaction and peer income becomes more significant as the peer group becomes more similar to the respondent. The study on the relationship between happiness and relative income position is not a new undertaking, and although recent studies, such as the ones conducted by Luttmer (2005) and Ferrer-i-Carbonell (2005) identify an inverse relation between happiness and peer income, Diener *et al*.(1995) and Inglehart (1990) suggest the opposite may be true: that individuals have greater happiness when they are surrounded by wealthier individuals.

This paper investigates the possibility that a similar phenomenon holds true for student satisfaction with teaching. We investigate whether an individual's evaluation of his or her instructor is influenced by peer performance, and we examine which of the many definitions of peer performance is most influential in the student's evaluation. Although overwhelming evidence suggests that a student's evaluation of his or her instructor is related to their expected grade, we ask whether it is the relative or absolute grade that matters, or if both are important.

Two papers have recently addressed this issue. Isley and Singh (2005) and McPherson (2006) find some evidence that student satisfaction is inversely related to the student's own prior grade history, but we are unaware of any attempt to decipher the relationship between student satisfaction with teaching and the student's expected grade relative to the grades of their peers. If student satisfaction is relative, a student's evaluation of teaching may decline as peer group performance increases relative to one's own grade. Alternatively, students may conclude that increases in peer performance are the result of better teaching, and thus reward teachers by giving them higher evaluations.

We conclude that student satisfaction with instruction depends not only on one's own experience in the classroom, but on the experiences of one's peers. We show that student satisfaction is related to both one's own expected grade and the grades of others around them. We investigate different measures of how the reference group may be determined, and conclude that individual students reward teachers with higher evaluations as both their own grade and the grades of their peers increase.

Background

Researchers have examined how instructor grades influence student evaluations of teaching (SET) for almost 75 years. Mason *et al*. (1995) noted that by the year 1990, over 1,300 articles had been published trying to explain how students evaluate faculty. This vast quantity of research reflects the fact that, despite wide criticism regarding the reliability of such instruments, student evaluations play an important role in the tenure and promotion process at most US colleges and universities. The instruments used for the evaluation of faculty are typically completed by students at the end of the semester, and require students to rate their instructors on a Likert scale indicating the perceived quality of instruction.

One of the most intriguing items in this literature is the relationship between expected student grades and SET ratings. Krautmann and Sander (1999) conclude (p. 61) that 'faculty have the ability to "buy"' higher [teaching] evaluations by lowering their grading standards.' By lowering grading standards, and making it less difficult for students to earn high grades, faculty can prompt students to respond with higher SET ratings. McPherson (2006, p. 18) comes to a similar conclusion when he claims that 'higher expected grades do lead to significantly better SET scores among both principles and upper division [economics] classes.'

A caveat to this finding was brought forth by Isley and Singh (2005) who found an inverse relationship between SET scores and historical grade point averages. They suggest that it is a student's expected grade relative to their customary grade that is the best predictor of SET ratings. Although all three of these papers provide interesting insights to the student evaluation process, these studies are based on class level data and, as a result, are not able to provide a robust test of how individual relative grade expectations may influence SET ratings.

Knowing whether relative grade expectations impact SET ratings is important because if students reward teachers for high relative grades as opposed to simply high absolute grades there may be limits to an instructor's ability to 'purchase' better teaching evaluations by increasing the grades of all students. Conversely, if individual students reward teachers for their own high grades as well as the high grades of their peers, it becomes expensive to give low grades to anyone in class and increases the incentive to 'buy' higher SET ratings. To be sure, the incentive depends on (a) how relative grade expectations are formed, and (b) who individual students consider as their peers. In this paper we choose to focus on how alternative measures of relative grades influences SET ratings.

We posit three different ways that students may identify relative peer groups and discuss the implications for SET ratings in each case. First, satisfaction with one's own grade may be referenced by a fixed point such as one's own historical GPA. When a student receives a grade higher than their fixed reference point they are relatively pleased with their instructor, and when a student earns a grade lower than their fixed reference point they are relatively disappointed and are less satisfied with their instructor. In this case, grades received by peers are not relevant, and the student's only reference point is his or her own historical performance. As such, if higher relative grades lead to higher student evaluations, it is possible for teachers to 'buy' higher evaluations simply by giving higher grades. This is the form of relative performance investigated by Isley and Singh (2005) when they used expected average class grade relative to average class cumulative GPA to help explain SET ratings. Although their study provides critical insight as to the importance of relative grade expectations on SET scores, their approach cannot answer questions regarding the impact of peer performance on SET ratings.

Second, satisfaction with one's own grade may be referenced by an aggregate measure of peer performance such as the average grade earned by all students who take the same course, or by the average grade given by the instructor during past semesters. Student satisfaction may be positively or negatively related to the grades earned by their peer group. If students are only concerned about their relative position in grading, they may penalise teachers for giving higher grades to their peers. Alternatively, students may interpret high grades among their peers as a sign of good teaching, and prompt them to further reward their teacher with high evaluations.

Third, students may use a reference point such as the perceived average grade in the individual class in which they are enrolled. Comparing one's own performance to the average class performance appears to be the norm. We have observed that most teachers, when they hand back a test, report the average or median class grade. Students often seem to take some relief in the fact that they have done better than average even if their performance is not absolutely satisfactory. Students also seem to believe that something has gone amiss when the average score reported by their teacher is lower than what they consider 'normal.' This suggests that students may react either positively or negatively to increases in the grades of their peers. Therefore, it seems reasonable to assume that students consider both their relative grade as well as the overall performance of the class.

In this third case, the students' subjective reference point is not fixed, but rather is determined by the grades given in the class of which they are part. As an instructor increases individual grades he or she also raises the average class grade. The increase in the average grade likely has two opposite effects: it lowers the relative satisfaction of all students in class, but increases the perceived ability of the teacher. Increases in peer performance may increase or decrease SET ratings.

Econometric model

It is common to posit that a student's evaluation of his or her teacher depends on class characteristics, the student's demographic characteristics, instructor characteristics, and the grade earned in the class by the student. Based on this formulation we could model the satisfaction of student i in class j as $S_{ij} = F(C, T, D, G)$, where S_{ij} represents the level of satisfaction of the student i in class j, C represents a vector of class characteristics, T represents a vector of instructor behaviours, D represents student demographic characteristics and effort, and G represents the student's expected grade in the class.

We propose an alternative model of student satisfaction that recognises that a student's level of satisfaction with his or her instructor may depend on the student's relative grade in the class as well as his or her absolute grade. Let G represent a measure of the perceived performance of a reference group. Now we may modify the above model by writing $S_{ij} = F(C, T, D, G, G/\underline{G})$.

Note that if satisfaction is linear in perceived relative grade and independent of the other arguments in the satisfaction function, then $G/\underline{G}$, this will not alter the average level of satisfaction with instruction in aggregate. This points out why using class level data to test whether relative grade perceptions influence student evaluations may be fruitless. If students base their relative performance on the performance of the class on average, this variable may not appear to influence SET ratings when class level data is used.

In order to test the influence of relative expected grades on SET ratings we use five different models, representing the different ways relative performance is considered. In each case we correct for the sample selection identified by Becker and Powers (2001) and insure that expected grades are truly exogenous using a Hausman specification test.

Survey and data

During the last week of the fall semester 2003 we asked students at a large public university in the US for information about the grade they expected to receive in a class. The survey was conducted in 32 separate courses, representing every class offered by the economics department during the semester. It was given immediately after the students completed instructor/course evaluations, and students were told the survey was voluntary and were assured their responses would remain anonymous. Even though we asked students to provide their student identification number (which clearly meant that, in spite of the prior mentioned assurances to the contrary, the anonymity of their written evaluations could be compromised), response rates for those present in the classrooms at the time of the survey were greater than 95 per cent. A few students declined to fill out the questionnaire, but less than 3 per cent of the students who actually filled out the questionnaire omitted their student identification number. Thus, of a potential enrolled student population of 1,016 in the courses surveyed, we have complete data on 716 students. This 70 per cent response rate suggests an absenteeism rate of about 25 per cent on the day(s) of the survey.

The instructor evaluation form asked students about five dimensions of instructor quality: organisation, willingness to respond to students, availability, respect for students, and overall contribution of the instructor. Students ranked their instructors on a scale of 1(low) to 7 (high) in each category. Similar to McPherson (2006) we used the average of the five responses (SET) as the measure of student satisfaction with their instructor.[1] We did estimate the models using only the responses to the question regarding the overall contribution of the instructor (with ordered probit), but found no substantial differences with averaging the SET ratings. After obtaining student responses, we surveyed faculty about the different grading practices used during the semester and obtained actual grades assigned to each student at the end of the semester.

Similar to Mason *et al*. (1995), we categorised our explanatory variables as teacher-related, class-related or student-related. Because our unit of observation was the individual, we were able to use more detailed information than if we had used class level data. Because we were interested in the influence of expected grade and expected grade relative to a peer group, we also included measures of these two variables. The data we used to estimate our model are given in Table 1.

We considered whether the instructor was part-time (ADJUNCT = 1) or full-time (ADJUNCT = 0), and we included a variable indicating the percentage of the student's grade that was based on testing (TESTP). We considered four class characteristics. We included a variable indicating whether the class was a lower division, introductory class, (LOWER = 1) or a more advanced, upper division class (LOWER = 0). We controlled for class size (SIZE) and for the number of times the class met each week (MEET). At the university where the study was conducted classes meet once, twice or three times per week. All classes that meet once per week are evening classes, so this variable reflects this fact. Because the classes surveyed include both economics courses and courses in quantitative analysis we included a variable indicating if the class is in economics (ECON = 1) or in quantitative analysis (ECON = 0). Finally, we include the demographic variables reflecting the student's age (AGE) and whether or not the student was male (MALE = 1) or female (MALE = 0). We also controlled for the student's self-reported level of effort in the class (EFFORT). Students reported a level of effort from 1 (low) to 7 (high).

Because our question of interest is how SET ratings are influenced by expected grades and expected grade relative to peer performance, we include variables for each student's expected grade (EXGRADE) and their expected grade relative to that of a reference point. We calculated three different measures of reference points based on the definition of peer performance and one reference point based on the student's own historical performance. We calculated the ratio of the student's

Table 1: Means and standard deviations

Variable	Mean	Standard deviation
	SET Equations n =716	
SET	5.5	1.2
ADJUNCT	.14	.35
TESTP	.56	.22
LOWER	.85	.35
SIZE	38.1	13.1
MEET	2.43	.70
ECON	.66	.47
AGE	22.8	4.19
MALE	.72	.45
EFFORT	3.9	.76
EXGRADE	3.04	.23
GPA	3.13	.55
AVSECTION	3.04	.23
AVCOURSE	3.04	.16
AVINST	3.04	.20
	Sample Selection Equation n = 1016	
COMPLETE	.71	.46
ADJUNCT	.14	.35
LOWER	.88	.33
SIZE	38.3	13.5
MEET	2.43	.73
ECON	.65	.48
AGE	22.7	4.22
MALE	.73	.45
GRADE	2.54	.20

expected grade to the average expected grade of students in the course section (AVESECTION), the average expected grade given in all sections of the course (AVECOURSE), and the average grade given by the instructor in all classes he or she teaches (AVEINST). We also include a measure of relative performance based on the student's past grade history (GPA). If one's satisfaction with instruction is more highly correlated with peer performance as the peer group becomes more closely related to one's self, we expect peer reference groups of the individual class (AVESECTION) to be more influential than the peer performance of all students taking the same subject (AVECOURSE) or the peer performance in classes taught by the same instructor (AVEINST).

It is worthwhile to note that students only have subjective estimates of peer performance. These estimates of peer performance are based on historical faculty reputations, informal discussion between students regarding mid-term examination scores, and publicly available SET evaluation scores. With certainty, student networks exist that provide information on which classes and instructors are most difficult, which instructors should be avoided, and which instructors should be sought after based on grading practices.

Estimation

Prior to estimating the determinants of SET rankings we turn to two econometric issues: endogeniety and sample selection. Endogeniety is of concern because of the inclusion of expected grade as a right-hand side variable. Higher student evaluations may result from better teaching, but we have no data on 'better teaching' to include in our estimation. Since higher grades may also be partly the result of 'better teaching' this variable may be correlated with the error term.

Evidence on the endogeniety between SET ratings and expected grades is mixed. Krautman and Sander (1999), Isley and Singh (2005) and McPherson (2006) find no evidence of endogeniety, whereas Nelson and Lynch (1984) suggest endogeniety in expected grade may be a problem. We used the regression test suggested by Woolridge (2006, p.532) and found endogeniety not to be an issue in our data set.[2]

The second econometric issue we confronted was sample selection. Because surveys were conducted at the end of the semester, the sample may be a good reflection of the end of term enrolment, but is unlikely to provide an accurate reflection of all students enrolled. In addition, students not in attendance on the day the sample was conducted might be significantly different from those in attendance and completing the survey. Using class level data, this sample selection is generally addressed by including class sizes as an explanatory variable: however, because we have individual data we estimate a sample selection model described by Heckman (1979).

Using data from all students who received grades in the class we estimated a sample selection equation predicting whether or not the student was present in class on the day evaluations were administered (COMPLETE = 1 if survey was completed and 0 otherwise). We estimated COMPLETE using the actual grade students received in class (GRADE), whether the class was taught by a part-time instructor (ADJUNCT), and whether the class was offered at the introductory level (LOWER). We controlled for class size (SIZE), and the number of times the class met each week (MEET). In addition we accounted for whether the class was in economics (ECON), the student's age (AGE) and the student's gender (MALE). In this equation we used actual grade rather than expected grade, because we were able to obtain actual grades for all students, regardless of whether they were present on the day the survey was given. Using a probit regression we estimated the equation COMPLETE = $\beta_0 + \beta_1$ADJUNCT + β_2LOWER + β_3SIZE + β_4MEET +β_5ECON + β_6AGE + β_7MALE + β_8GRADE + E, and saved λ, the inverse Mills ratio to correct for sample selection when estimating SET rankings.

Estimated results from the selection equation are presented in Table 2. Students who earned higher grades were more likely to be present in class on the day of the evaluations and complete the SET form, as were students enrolled in economics classes. Students in lower division classes were less likely to be present on the day teacher evaluations were conducted.

We estimated the five different models below to predict SET ratings and present the results for each model in Table 3.3

Table 2: Regression results: sample selection equation

Variable	Estimated coefficient	t-ratio
Constant	–.62	–1.41
ADJUNCT	–.07	–.37
LOWER	–.23	–1.39
SIZE	–.01	–1.21
MEET	.05	.53
ECON	.21	1.86*
AGE	.02	1.74*
MALE	–.11	–1.09
GRADE	.37	9.52**
n = 1016	Unrestricted LLF = –560.2 Restricted LLF = –617.4	P-Value=.001

** = significant at the .05 level in a two-tailed t-test
* = significant at the .10 level in a two-tailed t-test

$$SET_i = a + \Sigma B_i C_i + \Sigma \alpha_j T_i + \Sigma \theta_i D_i + \gamma_1 G_i + \delta_1 \lambda + E_i \quad (1)$$

$$SET_i = a + \Sigma B_i C_i + \Sigma\alpha_i T_i + \Sigma\theta_i D_i + \gamma_1 G_i + \gamma_2/GPA + \delta_1 \lambda + E_i \quad (2)$$

$$SET_i = a + \Sigma B_i C_i + \Sigma\alpha_i T_i + \Sigma\theta_i D_i + \gamma_1 G_i + \gamma_2 G_i/AVSECTION + \delta_1 \lambda + E_i \quad (3)$$

$$SET_i = a + \Sigma B_i C_i + \Sigma\alpha_i T_i + \Sigma\theta_i D_i + \gamma_1 G_i + \gamma_2 G_i/AVCOURSE + \delta_1 \lambda + E_i \quad (4)$$

$$SET_i = a + \Sigma B_i C_i + \Sigma\alpha_i T_i + \Sigma\theta_i D_i + \gamma_1 G_i + \gamma_2 Gi/AVINST + \delta_1 \lambda + E_i \quad (5)$$

In each of the equations, C_i represents class characteristics, T_i represents teacher characteristics, D_i represents student demographics and effort, G_i represents the students expected grade, and λ represents the inverse Mills ratio from the sample selection equation. The four different reference grades are G_i/GPA, G_i/AVSECTION, G_i/AVCOURSE, and G_i/AVINST.[4]

Results

Our findings are generally consistent with existing research, and are presented in Tables 3 and 4. Part-time faculty receive lower evaluations than do full-time faculty. Class size and instructor SET rating are inversely related. Instructors in economics classes tend to have higher SET ratings than instructors in quantitative analysis classes. We find that older students tend to give higher SET ratings and that students who put forth more effort in class also give higher SET ratings. Consistent with past research we find higher expected grades are correlated with higher SET ratings. Estimation of equation (2) confirms the findings of both Isley and Singh (2005) and McPherson (2006), who noted that as a student's expected grade increases relative to their historical GPA, they further reward their instructors with a higher SET ratings.

As shown in Table 3, the estimated coefficient on λ, the inverse Mills ratio, was not significant in any of the regression equations, indicating that sample selection is likely not an issue. Apparently, the students who were absent on the day evaluations were conducted were not significantly different from those who were present. Because of this we present ordinary least squares estimates in Table 4. The OLS estimates, and the estimates corrected for sample selectivity are similar, and we will focus our discussion on the OLS results due to the insignificance of λ in the sample selection equation.

Results presented in Table 4 indicate that peer performance does have significant influence on individual SET rankings in equations (3) and (5). Based on a likelihood ratio test comparing the log likelihood function in equation (1), which does not include a relative measure of expected grade, with the log likelihood functions of equation (3), which includes the explanatory variable (expected grade/average course section grade), or equation (5) which includes the explanatory variable (expected grade/average instructor grade) we reject the null hypothesis that

Table 3: Regression results: SET equations using sample selection

Variable	Model One		Model Two		Model Three		Model Four		Model Five	
	Estimated Coefficient	t-value	Estimated Coefficient	t-value	Estimated Coefficient	t-value	Estimated Coefficient	t-value	Estimated Coefficient	t-value
Constant	4.54**	7.26	4.61**	7.39	4.09**	6.25	4.56**	6.75	3.89**	5.90
ADJUNCT	−.61**	−3.24	−.61**	−3.72	−.45**	−2.26	−.61**	−3.18	−.40**	−2.03
TESTP	−.01**	−3.69	−.01**	−3.72	−.01**	−1.99	−.01**	−2.97	−.01*	−1.90
LOWER	.18	1.04	.20	1.15	.21	1.25	.17	.97	.24	1.40
SIZE	−.02**	−3.42	−.01**	−3.1	−.01**	−3.01	−.02**	−3.39	−.02**	−3.25
MEET	−.16*	−1.82	−.15*	−1.77	−.09	−.96	−.16*	−1.79	−.08	−.89
ECON	.48**	4.44	.44**	3.89	.48**	4.33	.48**	4.41	.47**	4.29
AGE	.02**	1.99	.02*	1.80	.02**	2.08	.02**	1.98	.02**	2.17
MALE	−.04	−.4	−.04	−.43	−.03	−.34	−.04	−.40	−.03	−.31
EFFORT	.23**	3.89	.23**	3.93	.23**	3.92	.23**	3.88	.24**	4.03
EXGRADE	.28**	3.73	.18*	1.76	.75**	3.12	.25	.63	1.11**	3.88
EXGRADE/GPA			.35*	1.73						
EXGRADE/ AVSECTION					−1.46**	−2.05				
EXGRADE/ AVCOURSE							.09	.07		
EXGRADE/ AVINST									−2.57**	−3.00
λ	−.29	−.84	−.52	−1.4	−.28	−.83	−.28	−.84		
Log Likelihood	−1081.0		−1079.2		−1078.3		−1080.4		−1075.9	

** = significant at the .05 level in a two-tailed t-test   * = significant at the .10 level in a two-tailed t-test

Table 4: Regression results: SET equations using OLS

Variable	Model One		Model Two		Model Three		Model Four		Model Five	
	Estimated Coefficient	t-value	Estimated Coefficient	t-value	Estimated Coefficient	t-value	Estimated Coefficient	t-value	Estimated Coefficient	t-value
Constant	4.26**	8.03	4.16**	7.76	3.81**	6.68	4.27**	7.28	3.60**	6.29
ADJUNCT	–.60**	–3.18	–.60**	–3.19	–.45**	–2.21	–.60**	–3.12	–.40**	–1.99
TESTP	–.01**	–3.71	–.01**	–3.75	–.01**	–2.01	–.01**	–2.99	–.01*	–1.89
LOWER	.13	.81	.13	.80	.17	1.07	.18	.13	.19	1.18
SIZE	–.01**	–3.42	–.02**	–3.40	–.01**	–3.16	–.02**	–3.39	–.02**	–3.25
MEET	–.16*	–1.80	–.15*	–1.80	–.09	–94	–.16*	–1.78	–.08	–.87
ECON	.51**	4.80	.5 0**	4.69	.50**	4.77	.51**	4.76	.49**	4.65
AGE	.02**	2.27	.02**	2.29	.02**	2.36	.02*	2.26	.03**	2.46
MALE	–.06	–.62	–.07	–.75	–.05	–.60	–.06	–.63	–.05	–.54
EFFORT	.23**	3.98	.24**	4.04	.24**	4.01	.24**	3.98	.24**	4.12
EXGRADE	.32**	5.64	.27*	3.89	.79**	3.35	.29	.73	1.16**	4.05
EXGRADE/GPA			.22	1.54						
EXGRADE/ AVSECTION					–1.46**	–2.04				
EXGRADE/ AVCOURSE							.09	.07		
EXGRADE/ AVINST									–2.56**	–2.98
Log Likelihood	–1087.4		–1086.3		–1085.2		–1087.4		–1082.9	

** = significant at the .05 level in a two-tailed t-test * = significant at the .10 level in a two-tailed t-test adding this new variable does not improve the predictive power of the model with a p-value of less than 5 per cent. We find no significant benefit to adding the variable (expected grade/GPA) in equation (2) or expected grade/average course grade) in equation (4). These results shed some light as to how the performance of peers impacts SET ratings.

The influence of expected grade and reference grade on SET ratings are given by $\partial SET/\partial EXGRADE = \gamma_1 + \gamma_2$ /(Reference Grade) and $\partial SET/\partial$(Reference Grade) = $-\gamma_2$[EXGRADE/(Reference Grade)^2]. In equation (2), where the reference grade is historical GPA, estimated values for γ_1 and γ_2 are both positive. Therefore, $\partial SET/\partial EXGRADE$ is positive, and $\partial SET/\partial$(Reference Grade) is negative, indicating that for the sample, as expected grade increases SET rankings increase, and as the historical GPA increases, SET ranking declines.

In equation (3), where the reference grade is the average grade received in the individual course section, and in equation (5), where the reference grade is the average grade given by the instructor, we find γ_1 to be significantly greater than zero, and γ_2 to be significantly less than zero. Therefore, $\partial SET/\partial EXGRADE$ may be positive or negative, and $\partial SET/\partial$(Reference Grade) is always positive. In these two cases it is possible that $\partial SET/\partial EXGRADE$ is negative, but only if $|\gamma_1| < |\gamma_2$ /(Reference Grade)$|$, which is not the case for our data.

The most striking aspect of our results is how different reference grades impact SET ratings. When a student's reference point is their own historical GPA, increases in their own expected grade have an unambiguously positive impact on their SET ratings. When the reference point is the expected grades of their peers in the same course section, or the grades of students who have taken the same instructor, increases in one's own expected grade may have a positive or negative impact on SET ratings, and increases in the reference grade have an unambiguously positive impact on their SET ratings.

Students positively reward the instructor for their own high grade, but this increase in the SET ranking is moderated by the fact that other students in class are not similarly rewarded. It is clear that peer group performance plays a significant role in how students evaluate their instructors, and it is clear that the definition of peer group matters. The peer groups that appear to have the largest influence on a student's SET rating are the students closest to the student, either sitting in the same class, or peers who have taken the same instructor.

Conclusion

It is commonly accepted that SET rankings are influenced by the expected grades, and that faculty are able to 'buy' higher SET ratings by giving higher grades.

Recently however researchers have questioned whether there are limits to the ability to buy grades due to the possibility that students reward teachers for their relative grade as opposed to their absolute grade. Some evidence exists that this may be the case. Using class level data researchers have identified an indirect relationship between average SET scores and average historical GPAs.

In this paper we use individual SET data to investigate the relationship between SET ratings and relative grades. Similar to the prior literature, we find an indirect relationship between SET scores and historical GPAs in our sample, but we find the opposite result to be true when we examine the relationship between SET scores and expected grade relative to average score given in the individual class and relative to the average grade given by the instructor. We have strong evidence that students reward instructors based on their own high grades, but also on the grades received by their peers. We find the most influential peer group to be those students sitting in the same class or the students taking the same instructor rather than the students enrolled in the same subject. Contrary to recent literature that suggests limits exist to an instructor's ability to purchase high SET scores when relative grades are considered, we find that the incentives to lower grading standards students and buy higher SET ratings may actually be greater than has been thought in the past.

Notes

1 By using the average SET rating across six categories we severely restrict the number of cases where SET equals either one or seven. In total, 29 of the 715 observations have an average SET value equal to either one or seven.

2 In order to conduct the endogeniety test we estimated the model EXGRADE = α_0 + α_1ADJUNCT + α_2TESTP + α_3LOWER + α_4SIZE + α_5MEET + α_6ECON + α_7AGE + α_8MALE + α_9EFFORT + $\alpha 10$ HSGPA + $\alpha 11$CRHOURS + α_{12}WORKHRS + V, where HSGPA represents a students High School grade point average, CRHOURS represents the number of credit hours taken by the student during the semester and WORKHRS represents the number of hours worked each week by the student. All other variables are defined above. We save the estimated error, V, and used the error as an explanatory variable in the estimation of SET. The second equation estimated is SET = α_0 + α_1ADJUNCT + α_2TESTP + α_3LOWER + α_4SIZE + α_5MEET + α_6ECON + α_7AGE + α_8MALE + α_9EFFORT + α_{10} EXGRADE+ α_{11}V + U, where U is the estimated error term. The instrumental variables used in the first equation are HSGPA, CRHOURS, and WORKHRS. HSGPA was positively related to EXGRADE, and both CRHRS and WORKHRS were negatively related to EXGRADE. Both the estimated coefficients on HSGPA and WORKHOURS were significant at α <.02. To test for endogeniety, we use the SET equation, and test Ho: $\alpha_{11} = 0$ against the two sided alternative. Based on a t-value of –1.36, we fail to reject the null hypothesis and conclude endogeniety is not an issue.

3 Initially, we estimated the equation $SET_i = a + \Sigma B_i C_i + \Sigma \alpha_j T_i + \Sigma \theta_i D_i + \gamma_1 G_i + \gamma_2 G_i/GPA + \gamma_3 G_i/AVSECTION + \gamma_4 G_i/AVCOURSE + \gamma_5 G_i/AVINST + \delta_1\lambda + E_i$, using all four different reference grades in the same equation. Although jointly these four

variables add significant explanatory power to the estimation of SET, multicollinearity between these four explanatory variables resulted in insignificant ratios for these variables as well as for the estimated coefficient on expected grade, Gi.

4 The inclusion of explanatory variables in the equations predicting SET that are not present in the selection equation is unorthodox and may result in inconsistent estimates. We believe the benefit of including data gathered in the SET process outweighs the disadvantage caused by the inconsistent estimates

References

Becker, W.E. and Powers, J. (2001) Student Performance, Attrition, and Class Size Given Missing Student Data. *Economics of Education Review*, 20, 377–388.

Diener, E., Diener, M. and Diener, C.L. (1995) Factors Predicting the Subjective Well-being of Nations. *Journal of Personality and Social Psychology*, 69, 851–864.

Duesenberry, J. (1949) *Income, Savings and the Theory of Consumer Behavior*. Cambridge, MA: Harvard University Press.

Ferrer-i-Carbonell, A. (2005) Income and Well-being: an Empirical Analysis of the Comparison Income Effect. *Journal of Public Economics*, 89, 997–1019.

Heckman, J.J. (1979) Sample Selection Bias as a Specification Error. *Econometrics*, 47, 153–161.

Inglehart, R. F. (1990) *Culture Shift in Advanced Industrial Society*. NJ: Princeton University Press.

Isley, P. and Singh, H. (2005) Do higher grades lead to favorable student evaluations? *Journal of Economic Education*, 36, 29–42.

Krautmann, A. C. and Sander, W. (1999) Grades and Student Evaluation of Teachers. *Economics of Education Review*, 18, 59–63.

Luttmer, E. F. (2005) Neighbors as Negatives: Relative Earnings and Well-being. Quarterly *Journal of Economics*, 102, 963–1002.

Mason, P. M., J. W. Steagall, and Fabritius, M. M. (1995) Student evaluations of Faculty: A New Procedure for using Aggregate Measures of Performance. *Economics of Education Review*, 14, 403–416.

McPherson, M.A. (2006) Determinants of How Students Evaluate Teachers. *Journal of Economic Education*, 37, 3–20.

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

Veblen, T. (1899) *The Theory of the Leisure Class*. New York: Macmillan.

Woolridge, J.M. (2006) *Introductory Econometrics, Third Edition*. Mason, Ohio: Thomson Southwestern.

Contact details

Professor Clifford Nowell

Department of Economics

Weber State University

Ogden UT 84403

Email: cnowell@weber.edu