Determinants of Student Evaluation of Teaching: Evidence from Turkey

Z. Bilgen Susanlı
Assistant Professor
Department of Economics, Işık University
Istanbul, Turkey
Email: susanli@isikun.edu.tr

Mehmet Kaytaz
Professor
Department of Economics, Işık University
Istanbul, Turkey
Email: kaytaz@isikun.edu.tr

Abstract

This paper analyzes the relationship between expected grades and student evaluations of teaching in the Faculty of Economics and Administrative Sciences in Işık University from Fall 2007 to Spring 2013. Findings indicate that there is a positive and significant association between student evaluations of teaching and expected grades. This result is robust to controlling for instructor and course fixed effects. Results using student-level data also suggest that student evaluations of teaching are negatively associated with the variance of the expected grades within a class. This is consistent with the hypothesis that students are risk averse and prefer less uncertainty about their expected course grades.

Keywords: Student evaluations of teaching, expected grades, fixed effects.

1. Introduction

It has been well-documented that college grade point averages in the United States rose substantially between the 1960s and the 2000s, which lead to the definition of *grade inflation* concept. This gave rise to a large body of research that studied the causes and consequences of grade inflation (Jewell, McPherson, & Tieslau, 2013; Kuh & Hu, 1999; Love & Kotchen, 2010). Theoretically, grade inflation may be an outcome of two factors. First, grades may simply rise due to increasing student quality and teaching quality/effort. Second, it may be an outcome of instructors' effort to improve their scores in the students' evaluations of their teaching.[1]

Much of the existing research investigates whether grade leniency does affect student evaluations of teaching (SET), which, in turn, puts the reliability of SETs into question. The reliability of SET is important for three reasons. One is the issue of measuring the quality of instruction. Considering that students' evaluations play a determining role for faculty evaluations procedures for tenure, promotion and pay schemes as well as applications for new faculty positions at different institutions, the reliability of SET as a measure of quality of instruction becomes questionable if evaluations can be improved by giving higher grades. Second, grade point average (GPA) in the undergraduate degree is one of the parameters that factor in the decisions of recruiters in the job market for the fresh graduates. Hence, grades may act as a signal in the job market, especially for the first-time applicants, and grade inflation may dilute this signal by compressing grades at the upper end of the distribution. In addition, undergraduate GPA is an important criterion in the graduate school admissions process. This becomes even more so when a student is applying to a program in another country where the admissions committee has limited knowledge about the sending institution. Third, it may also be argued that in response to inflated grades

[1] Love & Kotchen (2010) show that grade inflation arises as a result of institutional expectations placed on faculty members. Increased emphasis on course evaluations as well as research productivity may induce faculty to lower their teaching effort and to inflate grades.

students may lower their effort and invest less time in studying, which may lead to changes in human capital formation (Babcock, 2010).[2]

Numerous studies have explored the relationship between expected grades and SETs using data from the universities in the United States. These are built on the positive correlation between students' expected grades in a course and their evaluations for the instructor. Much of the empirical evidence is suggestive of the positive association between SETs and students' expected grades (Ewing, 2012; Krautmann & Sander, 1999; Langbein, 2008). Several variables are used as measures of grades in previous studies. While some authors use actual grades given by the instructor, some others use expected grades reported by students on the SET survey. Since students do not know their actual grades at the time of the SET surveys and report what grade they expect, some studies exploit the difference between the expected and actual grades as a determinant of how students rate instructors (Isely & Singh, 2005). Along similar lines, "relative expected grade", the difference between what grade students expect from the course and what grade they are accustomed to receive, is found to be a significant determinant of student evaluations (Ewing, 2012). In addition to expected grades, the distribution of expected grades across students may be a significant factor in determining evaluations. Student ratings are found to be negatively associated with the variance of expected grades suggesting that students are risk averse and hence they prefer less uncertainty about their grades (Matos-Díaz & Ragan, 2010).

This paper is motivated by the fact that contrary to what was observed in U.S. universities, GPAs in Işık University's Faculty of Economics and Administrative Science (FEAS) programs displayed a substantial decline between 2007 and 2013. The average GPA fell from 2.8 to 2.6 (out of 4 points) between 2007 and 2013 (Figure 1). A sharper trend is observed among the students of Economics and Political Science departments. While the average GPA of a student graduating from Economics program fell from 3.3 to 2.8 between 2007 and 2013, the GPAs of students in Management department was roughly stable around 2.6 throughout the same period (Figure 2). Moreover, in line with the aggregate trend the average grades received by students enrolled in courses given by programs under the FEAS also experienced a substantial decline from 2.9 in 2007 to 2.6 in 2013 (not presented). Then, we ask if, as suggested by the evidence from the U.S. universities, SETs are the driving factor behind grade inflation, how does the relationship between expected grades and SETs play out during a period of falling course grades ?

This paper analyzes the relationship between expected grades and SETs using both student-level and course-level data from SET of Işık University's FEAS programs. We explore the variation in expected grades across departments and over time. In order to take into account unobserved characteristics, we control for instructor, course and instructor-course fixed effects.

Our dataset provides a good laboratory for studying this topic since SETs factor into neither the pay system nor promotion decisions of the instructors at Işık University. Therefore, it is plausible to assume that there is very little, if any, incentive for instructors to "buy" higher evaluation scores by inflating grades. Moreover, the period of this study is characterized by falling GPAs, which further supports the validity of this assumption.

The contributions of this paper are twofold. First, it is the first study to examine the relationship between SETs and expected grades in an environment where, due to institutional standards, there is (almost) no incentive for instructors to inflate grades for getting higher SET. Second, to the best of our knowledge, this paper is the first to investigate the determinants of SET in the context of a Turkish university. Hence, they provide important insights for the university administrations and faculty members as well as recruiters and policymakers.

We find that there is a positive and significant association between the expected grades and the SETs. This finding is robust to controlling for instructor, course and instructor-course fixed effects. This result is stronger in courses taught by the department of Management. Results using student-level data indicate that SETs depend also on the distribution of expected grades within a class, which implies that students are risk averse and prefer less uncertainty about their grades.

The remainder of this paper proceeds as follows: Section 2 sets the background of the study and provides information about Işık University. Section 3 describes the dataset and the empirical methodology. Section 4 presents the results and the sensitivity analyses, and Section 5 concludes.

[2] Babcock and Marks (2011) present suggestive evidence for the fall in aggregate time spent studying by college students between 1960 and 2004. They show that average time spent studying declined from about 24 hours a week in 1960 to 14 hours a week in 2004.

2. Background

This study is conducted using both course-level and student-level data from the SETs administered at the Faculty of Economics and Administrative Sciences at Işık University. Işık University is a private (foundation), Ph.D.-granting institution located in Istanbul, Turkey. There are four faculties and two institutes with about 4,000 undergraduate students. The FEAS consists of three departments (Economics, Management and International Relations and Political Science), which provide seven undergraduate and six graduate programs. The language of instruction is English at both undergraduate and graduate level.

For the purposes of this study, we limit our sample to undergraduate courses since the dynamics governing the incentives and the satisfaction of graduate students may substantially be different.

It is important to note that compared to the system in the US, SETs play a much less critical role for assessing the performance of instructors in the Turkish academia partly because tenure system does not exist in the Turkish system. While SETs bear much less importance for the performance evaluation of instructors at Turkish universities, they are expected not to be at the extremes. In addition, due to the rise in the number of Turkish academics with doctoral degrees from U.S. institutions as well as the proliferation of foundation universities in Turkey, the academic job market in Turkey has started to resemble that in the US in the sense that when applying for a faculty position, an instructor may be asked to submit previous SETs. Therefore, although student evaluations are not critical to performance and promotion, they may certainly play a role if and when an instructor is switching to another institution.

3. Data and Empirical Methodology

We use course-level and student-level data from SET on FEAS courses during 2007-2013. The data was obtained from the Registrar's Office of Işık University and covers 738 undergraduate classes taught by 73 different instructors. In addition, we explore student-level data for 127 courses taught in Fall 2012 and Spring 2013.

A SET questionnaire is administered for each course in the last month of every semester. At the time of the survey the instructor leaves the classroom, and a graduate assistant administers the survey and hands the completed surveys to the Registrar's Office in order to ensure strict confidentiality of the students' replies. Figure 3 in the Appendix provides the survey questionnaire handed out to the students that are present in class at the time of the survey. No information regarding the identification of the students is collected in the survey. Hence, all forms are anonymous. When completing the survey, students are asked to rate several aspects of the instructor and the course as well as their own involvement and effort for that course during the semester. Of particular interest to us, are the overall ratings students give the instructors and the grade students expect. Table 1 reports the summary statistics of the variables used in this study. In our sample of 738 courses taught during 2007-2013, students rate their instructors on average 4.28 out of 5 when they are asked "*Overall, I would rate the instructor as: 1:Poor, 2:Fair, 3:Satisfactory, 4:Good, 5:Excellent"*. The average expected grade is 2.96 while the (actual) average grade received by all enrolled students is 2.34. It should be noted that while the expected grade variable is obtained from the responses from the students that were present at the time of the survey, the actual grades variable is calculated using grade data from all enrolled students after grades are submitted by instructors. This difference may also reflect overconfident grade expectations (Nowell & Alston, 2007). Summary statistics indicate that 43 percent of courses were taught by female instructors, 16 percent of them were taught by full professors and 74 percent were taught by full-time faculty members.

3.1 Basic OLS Model

Theoretically, students' evaluation of teaching is determined by several variables such as class size, the level of the course and the response rate, which is the share of enrolled students that filled the questionnaire. In addition, the instructor's gender, seniority, part-time full-time status as well as teaching productivity (or talent in teaching) can also be included as a determinant of SET. Since teaching productivity is not observed and if it is also correlated with students' expected grades then the ordinary least squares analyses would yield biased estimates.

Following Ewing (2012), we estimate the following equation as our baseline model:

$$\mathrm{SET}_j = \alpha_0 + \alpha_1 \mathrm{X}_j + \alpha_2 \mathrm{ExpGrade}_j + \varepsilon_j \qquad (1)$$

where SET_j is the outcome variable for the course *j*, calculated as the average of the students' rating of the instructor for course *j*, X_j vector of control variables, and ε_j is the error term [3]. The control variables captured by the vector X_j are class size measured by the number of enrolled students, total replies, the department that the course belongs to, and a dummy variable that equals one if the course is mandatory. We control also for instructor characteristics including instructor's gender, part-time full-time status and a dummy variable for whether the instructor is a full-professor.[4] Other control variables are the semester of the course, a dummy variable for spring semester as well as time dummies for calendar year. Our coefficient of interest α_2 captures the association between expected grades, ExpGrade, and SET. [5]

3.2 Sensitivity Analyses

Matos-Díaz and Ragan (2010) suggest that when studying the relationship between expected grades and SET, the distribution of expected grades should be accounted for in addition to the average of expected grades. On the assumption that students prefer knowing their grades with greater certainty, a narrower distribution of grades across students would imply reduced uncertainty.

Following Matos-Díaz and Ragan (2010), we estimate the following equation:

$$SET_j = \beta_0 + \beta_1 X_j + \beta_2 ExpGrade_j + \beta_3 VAR(ExpGrade_j) + \vartheta_j \quad (2)$$

where all variables are as in equation (1), $VAR(ExpGrade_j)$ is the variance of the expected grade across students that complete a SET for a given class estimated using student-level data from Fall 2012 and Spring 2013, and ϑ_j is the error term.

If students are categorized into two groups as strong and weak, then it would be the weak students that put a higher value on a compressed distribution of expected grades. Strong students may prefer to be differentiated from the weaker ones.[6] Hence, the effect of a lower variance in the distribution of expected grades is ambiguous as the negative effect of a larger variance may diminish as the average expected grade rises. To test this, in one specification we allow the impact of the grade distribution to vary with student quality, which may be measured by two alternative variables average actual grades and average expected grade of the class.

3.3 Unobserved Heterogeneity

Admittedly, expected grades are potentially endogenous. One could argue that if more productive instructors raise the grade expectations of students and are higher rated by students, not controlling for instructors' teaching productivity may lead to biased OLS estimates. In order to address this issue, we estimate a fixed effects model. Based on the assumption of time invariant unobserved instructor characteristics, instructor fixed effects capture instructors' characteristics that persist over time. This allows us to account for unobserved instructor characteristics that may be correlated with the expected grades and the SET. Controlling for instructor fixed effects also allows us to take into account the possibility of nonrandom allocation of instructors across courses (Ragan & Walia, 2010).

One could also argue that controlling only for instructor fixed effects makes the assumption that an instructor's teaching productivity is constant across all the courses he or she teaches and over time (Ewing, 2012). To the contrary, teaching productivity may differ in courses taught by the same instructor and may improve over time. To address this, we control for instructor-course fixed effects, which will account for characteristics that are specific to a course taught by a given instructor and that potentially affect SETs. In order to account for the unobserved course specific characteristics that may be correlated with the expected grades, we control for course fixed effects. Finally, we control for year fixed effects to capture any unobserved effects that are specific to a year.

[3] We also explore an alternative variable that is whether students recommend the instructor to others. Our results are qualitatively the same.

[4] While it would be informative to control for the composition of students, that information is not available in our dataset.

[5] Isely and Singh (2005) and Ewing (2012) construct Relative Expected Grade (REG) variable by subtracting cumulative GPA from the expected grade and use REG as their variable of interest. We are unable to do so due to data limitations.

[6] Asassfeh, Al-Ebous, Khwaileh, and Al-Zoubi (2014) document that students with low GPAs base their evaluations on nonacademic criteria including lenient grading.

Following McPherson (2006), we estimate the following equation:

$$SET_{ijt} = \gamma_0 + \gamma_1 X_{ijt} + \gamma_2 ExpGrade_{ijt} + \mu_i + \eta_j + \tau_t + \chi_{ijt} \quad (3)$$

where μ_i is the fixed effect specific to instructor *i*, η_j is the course-specific fixed effect, τ_t is the time fixed effect, χ_{ijt} is the error term, and all other variables are the same as described in equation (1). Again, our coefficient of interest γ_2 captures the effect of the expected grades on SET scores.

4. Results

4.1 OLS Results

Table 2 presents results from estimating equation (1). In column (1), we start by regressing the SET variable on dummy variables for departments, economics, political science with management as the base category and a dummy variable for spring semester as well as year dummies. The variable of interest has a positive and significant coefficient of 0.30. Results also show that while students rate an instructor of an economics course on average 0.20 points higher than the instructor of a management course, we find no significant effect for instructors of courses taught by the political science department. Next, in column (2), we proceed by including additional course characteristics such as total replies, total registered students, a dummy variable that equals 1 if the course is mandatory, and the year the course is regularly taken, freshmen year as the base group. The coefficient of ExpGrade is still positive and statistically significant. In column (3), we continue by adding instructor characteristics. Results show that our coefficient of interest remains robust. While being a full professor is associated with significantly higher SETs scores, SETs are not significantly correlated with the instructors' gender and full-time status.

We also check if our results are uniform across courses from different departments. To do this, we estimate equation (1) with the full set of controls separately for each department. Results in columns (4) to (6) confirm that ExpGrade is positively and significantly associated with SET in all departments except political science. The coefficient of interest has the largest magnitude in the subsample of management courses. This finding may reflect the fact that the average GPA of students graduating from the management department is significantly lower than students from economics and political science majors.[7] Hence, the students of the management department may be more sensitive about their grades and rate instructors accordingly.

4.2 Sensitivity Analyses

In addition to the average expected grade, the distribution of expected grades across students within a class can also be a determinant of SETs. Following Matos-Díaz & Ragan (2010), we assume that students are risk averse regarding their grades and prefer a tighter distribution of grades. Therefore, we hypothesize that students reward instructors for a low variance of expected grades and rate them better on course evaluations. To address this, we explore student-level data for the two most recent semesters available at the Registrar's Office and estimate the variance of the expected grades within a class, Var(ExpGrade). We start by estimating equation (2) using the student-level data. Column (1) in table 3 presents the results. It should be noted here that in column (1) the interpretation is between students and not between classes because within a class students face the same instructor and the same classroom environment. We find that, on average, within a class a one-point rise in a student's expected grade is associated with 0.10 points higher rating of the instructor.

Next, we include the variance of ExpGrade within each course. Based on the findings shown in column (2), a one point increase in Var(ExpGrade) is associated with a 0.46 point lower score for SET. Accordingly, a one standard deviation increase (0.75) in Var(ExpGrade) reduces the predicted SET by 0.35 points. This suggests that students are indeed risk averse and prefer less uncertainty about their grades. In addition, the coefficient on ExpGrade falls from 0.10 to 0.06 which suggests that if Var(ExpGrade) variable is not controlled for, ExpGrade is capturing information about the risk preference of students. Next, we investigate whether the impact of risk preference of students depends on the level of expected grades. To do this, in column (3) we control for interaction of Var(ExpGrade) with ExpGrade. Results show that the negative association between SET and Var(ExpGrade) is

[7] A t-test of equal means between management students and others suggests that GPAs of students with management majors are significantly lower than others with a t-value 16.30 and 736 degrees of freedom.

robust to controlling for its interaction with ExpGrade, which implies that students are indeed risk averse and the impact does depend on what letter grade the student expects from the course. For example, a one standard deviation increase (0.75) in Var(ExpGrade) lowers the predicted value of SET by 0.41 when ExpGrade is 3.5 but by 0.36 when ExpGrade is 2.5. This shows that the students penalize intructors more harshly for greater uncertainty regarding their grades as their average expected grade increases.

Next, we take into consideration the possibility that the impact of Var(ExpGrade) may depend on the quality of students as proxied by actual grades. One can argue that students' preferences regarding the distribution of grades in a course may vary with the quality of students in a class (Matos-Diaz & Ragan, 2010). Strong students may prefer a wider distribution of grades because it allows for differentiation between the strong and weak types. Therefore, while the strong students prefer less uncertainty about their grades, they may as well want greater differentation. Hence, the impact of grade distribution is ambigous and depends on the average student quality in a class. To address this, we control for the actual grades and its interaction with the variance of expected grades in column (4). While the coefficient of actual grades variable is positive and statistically significant, its interaction with Var(ExpGrade) is imprecisely estimated. We find no evidence that the impact of Var(ExpGrade) depends on student quality while SETs are positively and strongly correlated with student quality.

For the comparability of results with the course-level regression results, we aggregate the student-level data to obtain class-level data. Table 4 summarizes the corresponding results. In column (1), we start with the baseline estimation. We find that the coefficient of ExpGrade is greater when the sample is restricted to the last two semesters of our dataset. This finding may be interpreted as an over time increase in students' sensitivity about their grades as GPAs are on the decline. In column (2), we introduce Var(ExpGrade) and find that there is no significant association between the SETs and the variance of expected grades. In column (3), we add the interaction variable, ExpGrade*Var(ExpGrade) to test if the relationship between SET and Var(ExpGrade) depends on ExpGrade. In column (4), we test for the possibility that the impact of Var(ExpGrade) depends on the student quality, measured by the actual grades. Results in column (4) indicate that the coefficient on Var(ExpGrade) and its interaction with the actual grades is not statististically significant. While the association between SET and expected grades remains robust throughout columns (2) – (4), we find no evidence for the association between the SETs and Var(ExpGrade) at the course level.

Results shown in tables 3 and 4 imply that while an average student rates an instructor lower when his or her grading scheme generates a wider distribution of grades (and hence greater unpredictability), across classes there is no significant correlation betweeen the SETs and the predictability of grades.

4.3 Addressing Unobserved Heterogeneity

To take into account unobserved instructor and course characteristics, we control for instructor and course fixed effects. Table 5 summarizes the results from estimating Equation (4).[8] Column (1) presents the results from controlling for only instructor fixed effects in the baseline regression. Controlling for instructor fixed effects holds constant the instructors' unobserved characteristics that are constant over time across classes. The coefficient of interest remains positive and significant. In column (2), course fixed effects are added in order to capture course characteristics that were initially unobserved. The magnitude of the coefficient on expected grades changes only slightly and remains significant. In column (3), instructor and instructor-course fixed effects are added in order to take into account the possibility that an instructor's teaching productivity may vary across the courses that he or she teaches.

These findings indicate that OLS results are robust to controlling for instructor and course fixed effects as well as instructor-course fixed effects. Even after taking into unobserved characteristics of the course and the instructor, expected grades seem to be the main determinant of students' assessment of teaching.

5. Conclusions

This paper aims to study the determinants of SET in an environment where there is little, if any, incentive for instructors to "buy" higher evaluation scores by inflating grades. To the contrary, over the period of this study overall GPAs declined by 0.20 points with substantial falls in GPAs in the two of three departments. While explaining the decline in GPAs in the FEAS is beyond the scope of this paper, we find robust evidence for the

[8] In fixed effects estimations, sample size falls to 683 since courses that are taught only once in the sample are excluded from the sample.

relationship between students' evaluations of teaching and the expected grades. Results suggest that holding all else constant, a one point increase in average expected grade leads to a 0.25 point (out of 5) increase in the evaluation score of an instructor.

Whether our results hold at other institutions require further analyses but these findings have important implications. First, our findings suggest that even after controlling for factors that are expected to be important for students' evaluations, expected grades still matter. Second, all else constant, student quality (proxied by actual grades) appears to be another important determinant of student evaluations, which highlights the importance of student admissions process for a university. Finally, the varying impact of expected grades on SETs across departments should be taken into consideration if and when teaching quality and effectiveness are assessed based on students' evaluations of teaching.

References

Asassfeh S., Al-Ebous, H. , Khwaileh, F., & Al-Zoubi, Z. (2014). Student faculty evaluation (SFE) at Jordanian universities: a student perspective. *Educational Studies*, 40(2), 121-143. doi: 10.1080/03055698.2013.833084

Babcock, P. (2010). Real costs of nominal grade inflation? New evidence from student course evaluations. *Economic Inquiry*, 48(4), 983–996. doi:10.1111/j.1465-7295.2009.00245.x

Babcock, P., & Marks, M. (2011). The falling time cost of college: evidence from half a century of time use data. *The Review of Economics and Statistics*, 93(2), 468–478.

Ewing, A. M. (2012). Estimating the impact of relative expected grade on student evaluations of teachers. *Economics of Education Review*, 31(1), 141–154. doi:10.1016/j.econedurev.2011.10.002

Isely,P. & Singh, H. (2005). Do higher grades lead to favorable student evaluations ?. *The Journal of Economic Education*, 36(1), 29–42.

Jewell, R. T., McPherson, M. A., & Tieslau, M. A. (2013). Whose fault is it? Assigning blame for grade inflation in higher education. *Applied Economics*, 45(9), 1185–1200. doi:10.1080/00036846.2011.621884

Krautmann, A., & Sander, W. (1999). Grades and student evaluations of teachers. *Economics of Education Review*, 18(4), 59–63. Retrieved from http://www.sciencedirect.com/science/article/pii/S0272775798000041

Kuh, G. D., & Hu, S. (1999). Unraveling the complexity of the increase in college grades from the mid-1980s to the mid-1990s. *Educational Evaluation and Policy Analysis*, 21(3), 297–320. doi:10.3102/01623737021003297

Langbein, L. (2008). Management by results: Student evaluation of faculty teaching and the mis-measurement of performance. *Economics of Education Review*, 27(4), 417–428. doi:10.1016/j.econedurev.2006.12.003

Love, D. A., & Kotchen, M.J. (2010). Grades, course evaluations, and academic incentives. *Eastern Economic Journal*, 36(2), 151–163. doi:10.1057/eej.2009.6

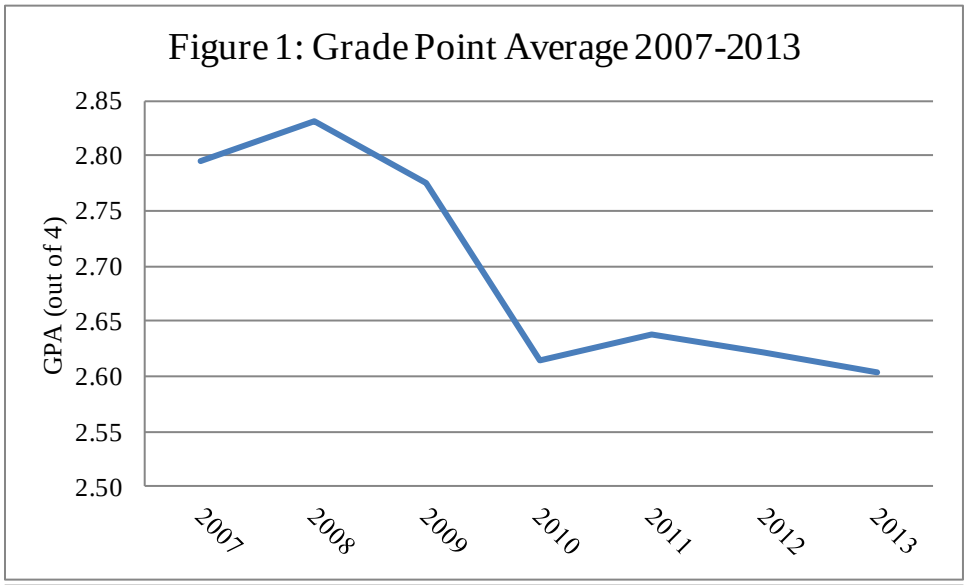
Matos-Díaz, H., & Ragan, J. F. (2010). Do student evaluations of teaching depend on the distribution of expected grade? *Education Economics*, 18(3), 317–330. doi:10.1080/09645290903109444

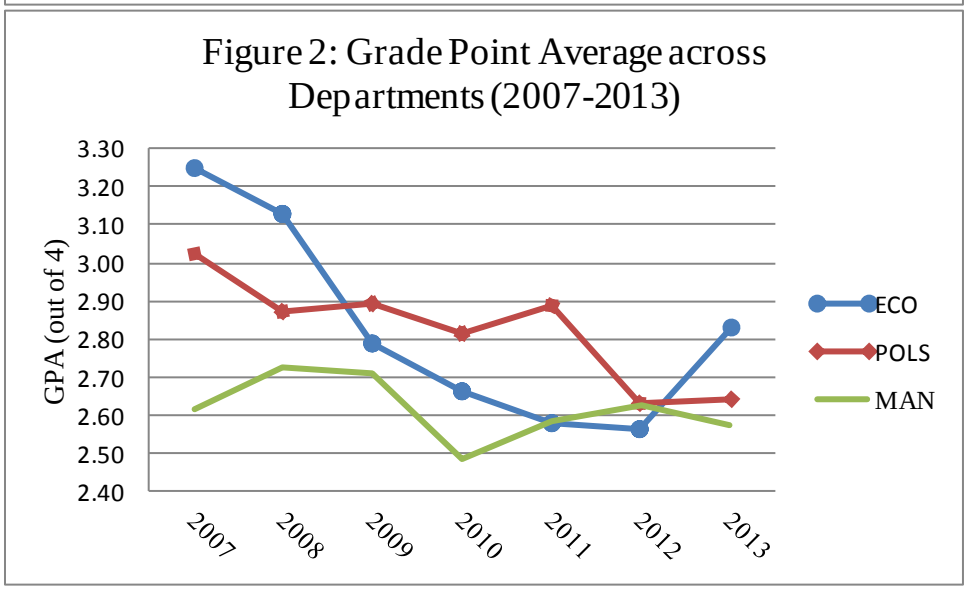
McPherson, M. A. (2006). Determinants of how students evaluate teachers. *The Journal of Economic Education*, 37(1), 3–20. doi:10.3200/JECE.37.1.3-20

Nowell, C., & Alston, R. M. (2007). I thought I got an A! Overconfidence across the economics curriculum. *The Journal of Economic Education*, 38(2), 131–142. doi:10.3200/JECE.38.2.131-142

Ragan, J. F., & Walia, B. (2010). Differences in Student Evaluations of Principles and Other Economics Courses and the Allocation of Faculty across Courses. *The Journal of Economic Education*, 41(4), 335–352. doi:10.1080/00220485.2010.510389

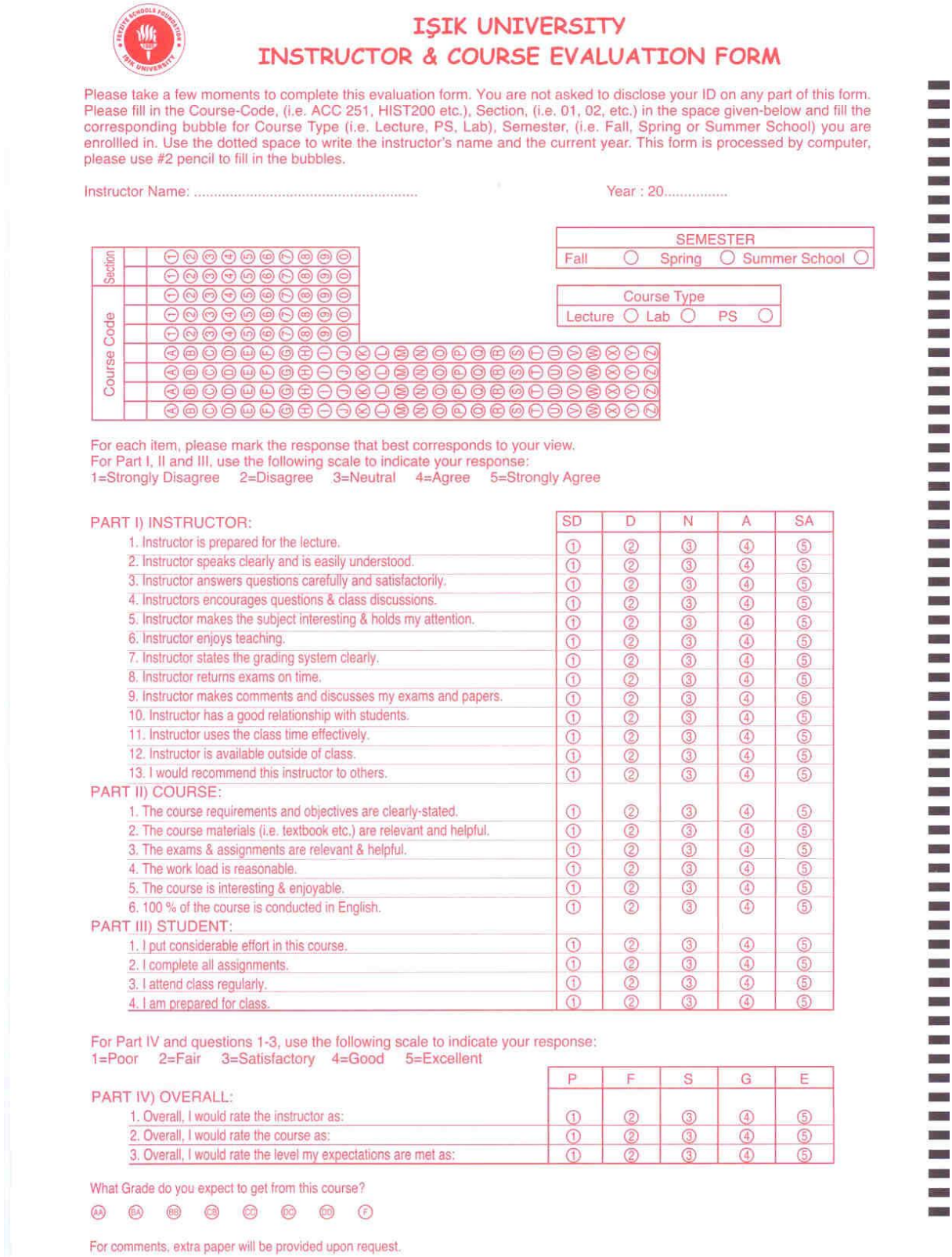
Appendix





Source: Authors' calculations using SET data from the Registrar's Office of Işık University.

Figure 3: Survey Questionnaire



IŞIK UNIVERSITY
INSTRUCTOR & COURSE EVALUATION FORM

Please take a few moments to complete this evaluation form. You are not asked to disclose your ID on any part of this form. Please fill in the Course-Code, (i.e. ACC 251, HIST200 etc.), Section, (i.e. 01, 02, etc.) in the space given-below and fill the corresponding bubble for Course Type (i.e. Lecture, PS, Lab), Semester, (i.e. Fall, Spring or Summer School) you are enrollled in. Use the dotted space to write the instructor's name and the current year. This form is processed by computer, please use #2 pencil to fill in the bubbles.

Instructor Name: ..... Year : 20.....

Section

Course Code

SEMESTER: Fall ○ Spring ○ Summer School ○

Course Type: Lecture ○ Lab ○ PS ○

For each item, please mark the response that best corresponds to your view.
For Part I, II and III, use the following scale to indicate your response:
1=Strongly Disagree 2=Disagree 3=Neutral 4=Agree 5=Strongly Agree

	SD	D	N	A	SA
PART I) INSTRUCTOR:					
1. Instructor is prepared for the lecture.	①	②	③	④	⑤
2. Instructor speaks clearly and is easily understood.	①	②	③	④	⑤
3. Instructor answers questions carefully and satisfactorily.	①	②	③	④	⑤
4. Instructors encourages questions & class discussions.	①	②	③	④	⑤
5. Instructor makes the subject interesting & holds my attention.	①	②	③	④	⑤
6. Instructor enjoys teaching.	①	②	③	④	⑤
7. Instructor states the grading system clearly.	①	②	③	④	⑤
8. Instructor returns exams on time.	①	②	③	④	⑤
9. Instructor makes comments and discusses my exams and papers.	①	②	③	④	⑤
10. Instructor has a good relationship with students.	①	②	③	④	⑤
11. Instructor uses the class time effectively.	①	②	③	④	⑤
12. Instructor is available outside of class.	①	②	③	④	⑤
13. I would recommend this instructor to others.	①	②	③	④	⑤
PART II) COURSE:					
1. The course requirements and objectives are clearly-stated.	①	②	③	④	⑤
2. The course materials (i.e. textbook etc.) are relevant and helpful.	①	②	③	④	⑤
3. The exams & assignments are relevant & helpful.	①	②	③	④	⑤
4. The work load is reasonable.	①	②	③	④	⑤
5. The course is interesting & enjoyable.	①	②	③	④	⑤
6. 100 % of the course is conducted in English.	①	②	③	④	⑤
PART III) STUDENT:					
1. I put considerable effort in this course.	①	②	③	④	⑤
2. I complete all assignments.	①	②	③	④	⑤
3. I attend class regularly.	①	②	③	④	⑤
4. I am prepared for class.	①	②	③	④	⑤

For Part IV and questions 1-3, use the following scale to indicate your response:
1=Poor 2=Fair 3=Satisfactory 4=Good 5=Excellent

	P	F	S	G	E
PART IV) OVERALL:					
1. Overall, I would rate the instructor as:	①	②	③	④	⑤
2. Overall, I would rate the course as:	①	②	③	④	⑤
3. Overall, I would rate the level my expectations are met as:	①	②	③	④	⑤

What Grade do you expect to get from this course?

AA BA BB CB CC DC DD F

For comments, extra paper will be provided upon request.

Table 1: Summary statistics of the variables used in the analyses.

Variable	Mean	Std. Dev.	Min	Max
Instructor's rating	4.28	0.54	1	5
Average Expected Grade	2.96	0.56	0	4.61
Class Average Grade (actual)	2.34	0.53	0.69	4
Total replies in class	13.59	9.89	1	55
Total registered students	34.97	19.51	2	104
Economics (ECO)	0.35	0.48	0	1
Management (MAN)	0.42	0.49	0	1
International Relations & Political Science (POLS)	0.23	0.42	0	1
Course is mandatory	0.62	0.49	0	1
Freshman year course	0.20	0.4	0	1
Sophomore year course	0.18	0.38	0	1
Junior year course	0.37	0.48	0	1
Senior year course	0.26	0.44	0	1
Instructor: Female	0.43	0.49	0	1
Instructor: Full Professor	0.16	0.36	0	1
Instructor: Full-time	0.74	0.43	0	1
Spring semester	0.54	0.5	0	1
Year:2007	0.11	0.32	0	1
Year:2008	0.13	0.33	0	1
Year:2009	0.15	0.35	0	1
Year:2010	0.17	0.38	0	1
Year:2011	0.17	0.38	0	1
Year:2012	0.17	0.38	0	1
Year:2013	0.09	0.29	0	1
Number of observations	738			

Source: Authors' calculations using data from Işık University Registrar's Office

Table 2: OLS regressions – using course-level data.

	(1)	(2)	(3)	(4)	(5)	(6)
Expected Grade (Avg)	0.30***	0.25***	0.26***	0.42***	0.18**	0.17
	[0.05]	[0.05]	[0.06]	[0.07]	[0.08]	[0.12]
Spring semester	-0.03	-0.03	-0.04	-0.07*	0.03	-0.04
	[0.04]	[0.04]	[0.04]	[0.04]	[0.06]	[0.09]
Total replies in class	-	-0.00	-0.00	0.00	-0.01	-0.01
		[0.00]	[0.00]	[0.00]	[0.01]	[0.01]
Total registered students	-	-0.00	-0.00	-0.00	-0.00	-0.00
		[0.00]	[0.00]	[0.00]	[0.00]	[0.01]
Mandatory course	-	-0.04	-0.06	-0.01	-0.03	-0.24**
		[0.06]	[0.07]	[0.16]	[0.09]	[0.10]
Sophomore year course	-	-0.02	-0.02	-0.17	0.04	0.07
		[0.10]	[0.10]	[0.28]	[0.08]	[0.20]
Junior year course	-	0.04	-0.02	0.02	-0.00	-0.17
		[0.08]	[0.06]	[0.11]	[0.11]	[0.17]
Senior year course	-	0.19**	0.16*	0.15	0.13	0.13
		[0.08]	[0.08]	[0.14]	[0.11]	[0.19]
Instructor Full-time	-	-	0.12	0.01	0.04	0.35**
			[0.10]	[0.16]	[0.12]	[0.16]
Instructor Full Prof	-	-	0.18	0.43***	0.06	0.00
			[0.12]	[0.15]	[0.14]	[0.00]
Instructor Female	-	-	-0.03	0.05	0.02	-0.01
			[0.07]	[0.10]	[0.14]	[0.11]
Economics	0.20**	0.22**	0.18*	-	-	-
	[0.09]	[0.09]	[0.10]			
Political Science	0.08	0.08	0.06	-	-	-
	[0.09]	[0.10]	[0.09]			
Departments included	All	All	All	MAN	ECO	POLS
Observations	738	737	737	312	255	170
R-squared	0.132	0.171	0.193	0.307	0.167	0.256
Adjusted R-squared	0.120	0.152	0.172	0.267	0.107	0.179

All columns estimate equation (1) and control for year fixed effects. Dependent variable is students' ratings of instructors. Robust standard errors clustered at the instructor level in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Sensitivity analyses – using student-level data.

Table 3: Sensitivity Analyses - Student-level regressions

	(1)	(2)	(3)	(4)
Expected Grade	0.10***	0.06***	0.11***	0.07***
	[0.02]	[0.01]	[0.02]	[0.02]
Total replies in class	-0.02***	-0.01***	-0.01**	-0.01**
	[0.01]	[0.00]	[0.00]	[0.00]
Total registered students	-0.00	-0.00	-0.00	-0.00*
	[0.00]	[0.00]	[0.00]	[0.00]
Economics	-0.00	0.00	0.01	-0.00
	[0.10]	[0.09]	[0.09]	[0.09]
International Relations	-0.07	0.08	0.08	0.01
	[0.15]	[0.13]	[0.13]	[0.13]
Mandatory course	0.02	0.03	0.03	0.05
	[0.07]	[0.07]	[0.07]	[0.06]
Sophomore year course	-0.00	-0.11	-0.10	-0.12
	[0.13]	[0.11]	[0.11]	[0.11]
Junior year course	0.23*	0.18	0.18	0.10
	[0.14]	[0.12]	[0.11]	[0.12]
Senior year course	0.33**	0.21*	0.21*	0.07
	[0.14]	[0.12]	[0.12]	[0.13]
Instructor Full-time	-0.34***	-0.33***	-0.32***	-0.31***
	[0.10]	[0.08]	[0.08]	[0.08]
Instructor Full Prof	0.22	0.14	0.14	0.14
	[0.13]	[0.13]	[0.13]	[0.11]
Instructor Female	-0.19	-0.19**	-0.19**	-0.16**
	[0.12]	[0.09]	[0.09]	[0.08]
Var(ExpGrade)	-	-0.46***	-0.31***	-0.06
		[0.08]	[0.10]	[0.26]
ExpGrade*Var(ExpGrade)	-	-	-0.05**	-0.02
			[0.02]	[0.02]
Actual Grades	-	-	-	0.30***
				[0.09]
Actual Grades*Var(ExpGrade)	-	-	-	-0.13
				[0.13]
Observations	1,493	1,492	1,492	1,492
R-squared	0.513	0.614	0.616	0.644
Adjusted R-squared	0.509	0.611	0.613	0.640

All columns estimate equation (1) and control for year fixed effects. Dependent variable is students' ratings of instructors. Robust standard errors clustered at the instructor level in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Sensitivity analyses – using course-level data.

	(1)	(2)	(3)	(4)
Expected Grade (Avg)	0.36**	0.49***	0.63***	0.55***
	[0.16]	[0.10]	[0.14]	[0.13]
Total replies in class	-0.01	-0.00	-0.00	-0.00
	[0.01]	[0.01]	[0.01]	[0.01]
Total registered students	-0.00	-0.00*	-0.00*	-0.00*
	[0.00]	[0.00]	[0.00]	[0.00]
Economics	0.15	0.10	0.13	0.12
	[0.10]	[0.09]	[0.09]	[0.10]
International Relations	0.04	0.12	0.11	0.09
	[0.12]	[0.11]	[0.11]	[0.12]
Mandatory course	0.01	0.06	0.05	0.06
	[0.07]	[0.07]	[0.07]	[0.06]
Sophomore year course	-0.04	-0.10	-0.07	-0.09
	[0.13]	[0.14]	[0.14]	[0.15]
Junior year course	0.24*	0.18	0.21*	0.16
	[0.13]	[0.12]	[0.12]	[0.13]
Senior year course	0.36**	0.22	0.23*	0.14
	[0.15]	[0.13]	[0.13]	[0.15]
Instructor Full-time	-0.19**	-0.24**	-0.19**	-0.20**
	[0.09]	[0.10]	[0.09]	[0.08]
Instructor Full Prof	0.01	0.00	-0.02	-0.03
	[0.15]	[0.13]	[0.13]	[0.12]
Instructor Female	-0.13	-0.12	-0.12	-0.10
	[0.10]	[0.09]	[0.09]	[0.08]
Var(ExpGrade)	-	-0.14	0.40	0.47
		[0.09]	[0.37]	[0.40]
ExpGrade*Var(ExpGrade)	-	-	-0.20	-0.14
			[0.13]	[0.13]
Actual Grades (Class Avg)	-	-	-	0.19*
				[0.11]
Actual Grades*Var(ExpGrade)	-	-	-	-0.11
				[0.10]
Observations	127	126	126	126
R-squared	0.527	0.636	0.646	0.656
Adjusted R-squared	0.472	0.590	0.598	0.602

All columns estimate equation (1) and control for year fixed effects. Dependent variable is students' ratings of instructors. Robust standard errors clustered at the instructor level in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Fixed effects regressions – using course-level data.

	(1)	(2)	(3)
Expected Grade (Avg)	0.23***	0.24***	0.23***
	[0.06]	[0.06]	[0.06]
Total replies in class	-0.01**	-0.00**	-0.00
	[0.00]	[0.00]	[0.00]
Total registered students	-0.00**	-0.00**	-0.01**
	[0.00]	[0.00]	[0.00]
Spring semester	-0.02	0.03	0.03
	[0.04]	[0.04]	[0.06]
Mandatory course	-0.08	-	-
	[0.06]		
Sophomore year course	-0.05	-	-
	[0.08]		
Junior year course	-0.15*	-	-
	[0.08]		
Senior year course	-0.05	-	-
	[0.08]		
Course fixed effects ?	No	Yes	No
Instructor-course fixed effects ?	No	No	Yes
Instructor fixed effects ?	Yes	Yes	Yes
Observations	683	683	683
R-squared	0.526	0.619	0.713
Adjusted R-squared	0.461	0.503	0.566

All columns estimate equation (1) and control for year fixed effects. Dependent variable is students' ratings of instructors. Robust standard errors clustered at the instructor level in brackets. *** p<0.01, ** p<0.05, * p<0.1