



The Movement to Online Course Evaluations: Do we hear from more complainers?¹

Elizabeth Breitbach^{a2}, Chandini Sankaran^b, Jamie Wagner^c

^a Department of Economics, University of South Carolina, United States, Elizabeth.Breitbach@moore.sc.edu

^b Department of Economics, University of South Carolina, United States

^c Department of Economics, University of Nebraska at Omaha, United States

Abstract

Student evaluations are an important source of feedback for instructors and college administrators. As many schools transition to online course evaluations, the question of whether the method of distribution affects the results arises. One reason the authors expect differences is due to the population each method of distribution reaches; the in-class distribution only allows students present on the day of the evaluation to participate, while the online evaluation reaches all students enrolled in the course. This paper will investigate the difference in response rates and make-up of who completes both the in-class and online evaluations for introductory economic courses over three semesters. Results indicate a two percentage point increase in participation on online evaluations leading to differences in aggregate student attributes and ratings for some important instructor performance variables. Upon further investigation, splitting the students into principles and introductory classes, show significant differences in instructor performance variables between the two methods of distribution.

Key Words: course evaluations, evaluation scores, student feedback, online evaluation

JEL Codes: A10, A20, A22

¹ The authors would like to thank Jason Beever, The Assessment Coordination Assistant at the Office of Institutional Research and Assessment at the Darla Moore School of Business for his valuable input and assistance. We would also like to thank the anonymous reviewer from PEER for their careful review and valuable comments to improve this paper.

² Corresponding author

1. Introduction

The purpose of student evaluations of teaching (SET) is to inform the instructor and college administrators of strengths within the course, while also offering constructive criticism. In-class evaluations have been the norm in many economics departments until recently, with some making the decision to transition to online distribution. In 2012, approximately 35 percent of economics departments administered SET online (Becker et al. 2012).

The move to online SET is beneficial for some important reasons. First, the online evaluations do not require the typical paper/pencil scantrons, removing resource costs associated with the SET. Second, valuable class time is not lost due to in-class evaluations; typically the online SET is completed outside the classroom at students' convenience. While the in-class SET can only be completed by students attending class on the date and time the evaluation is administered, the online SET gives access to all students enrolled in the class, even those not present on a specific day (Kronholm, et al. 1999; Dommeyer, et al. 2004). Though there are benefits to the online administration of SET, several important concerns arise. With more students having access to the online SET in a less structured environment, how will the response rates and the SET scores instructors receive be impacted?

The online SET has the potential to elicit higher response rates since all students have access to the evaluations, compared to only those in attendance on the date and time the in-class evaluation is administered. Students also have the ability to complete the evaluation any time of day and over a longer time period, further encouraging participation. Given the feeling of anonymity and more time, students may take the opportunity to be more candid with their feedback. However, while more students have the ability to respond to the online SET, the incentive has decreased. Apathetic students may not complete the online SET since the pressure of completing the survey in class is absent. The method of distribution could impact the SET response rates and transform the composition of students completing the evaluation; this in turn could affect instructor performance and course ratings.

The majority of past studies have found lower participation rates on the online evaluations, resulting in concerns from faculty about the accuracy of these results. Instructors are primarily appraised on their teaching effectiveness based on information obtained from SET. These results not only provide important feedback to individual instructors, but are also commonly required by most universities in tenure and promotion files (Becker et al. 2012). This research will aim to address the question of whether the method of distribution affects participation rates and mean SET scores, an important question for both instructors and administrators.

Data used in this study administered both in-class and online SET within sections of introductory and principles of economics courses. Comparisons will be made by differencing the aggregate student characteristics, response rates, and mean evaluation scores across methods of distribution, within the same section of the same course. Information on the response rates, aggregate self-reported student characteristics, and instructor performance and course ratings were collected and differenced across methods of distribution. An examination of the differenced values provides insight into the question of whether online distribution impacted the SET scores.

The aggregate results indicate few significant differences in mean course ratings across method of distribution in the individual SET questions. Results also suggest no significant difference in participation rates between online and in-class distribution methods for sections. However, the composition of students completing the SET has changed; there is a statistically larger portion of students who expect to fail completing the online SET. Students participating in-class gave higher evaluations to questions about the teacher preparations and knowledge, but lower scores to the textbook's helpfulness.

The division of general principles courses primarily enrolled by business and economics majors from introductory sections designed for business minors and other majors across the University tells a more interesting story. Although not all course rating results are significantly different between methods of distribution, the change in ratings for important indicators of teaching such as instructor preparedness, difficulty of exams and assignments, and class materials should be acknowledged. When the sample is divided by course type, a significantly higher response rate for the online SET is observed in introductory sections. Due to the increase in response rate, the question becomes has the "typical" student completing the evaluation changed? Our results suggest students expecting a B represent a greater portion of in-class evaluations; while those expecting to fail make up a larger portion of the online evaluation for introductory sections, both results are significant. There are few significant differences when looking at the specific questions asked on the SET.

2. Literature Review

Existing research has used various approaches to analyze the differences between ratings and response rates across method of SET distribution (Opengart and Mau 2012; Fike et al. 2010; Avery et al 2006; Donovan et al. 2006; Dommeyer et al 2004; Layne et al 1999). These studies test the difference between online and in-class SET scores and response rates, holding instructor characteristics constant. In order to achieve this, some studies compare different sections of a course taught by the same instructor (Stowell et al 2012; Opengart and Mau 2012; Donovan et al 2006; Dommeyer et al 2004) sometimes across semesters (Avery et al 2006). Other studies have randomized students who attended class on the day of the in-class evaluation into one mode of delivery or the other (Morrison 2013; Morrison 2011; Fike et al. 2010; Layne et al 1999). Lower response rates for the online evaluations were found using both types of instructor controls (Stowell 2012; Fike et al 2010; Avery et al 2006; Dommeyer et al. 2004; Donovan et al. 2006; Layne et al 1999).

Even though Avery et al. (2006) found a lower response rate with the web-based SET in their initial study, they provided evidence that online SET response rates will approach the levels of paper evaluations over time as students adjust to the new system. Furthermore, the differing response rates leads to the question of how the method alters the composition of students participating in the SET, and the impact this may have on SET scores.

Some studies have addressed the question of how the method of distribution has influenced participation by student characteristics (Kherfi 2011; Avery et al. 2006; Layne et al 1999). These studies have matched up official student data provided by the university's administration with their online responses on the SET. Students with strong opinions, those taking the course as part of their major, with heavier course loads, with higher GPAs, who

anticipate a higher final grade, those enrolled in smaller class sections, underclassmen, non-Asian students, and females are more likely to complete the online evaluations (Kherfi 2011; Avery et al. 2006; Layne et al 1999). The results imply that changing the composition of students participating in the SET may impact the mean evaluation scores.

Research has examined the impact method of distribution has on mean SET scores, finding mixed results. Some research concludes that the mode of delivery has no influence on the mean evaluation scores of instructors (Stowell 2012; Avery et al 2006; Donovan et al 2006; Dommeyer 2004; Layne et. al. 1999), others report higher scores for the in-class SET (Morrison 2013; Fike et. al. 2010; Opengart and Mau 2012; Hardy 2003; Ewell 2000), and still another finds that students who complete online respond more favorably (Carini et al 2003).

The data collected for this study improves on existing literature by comparing methods of distribution administered within the same section of the same course, eliminating bias from differences in section and student characteristics. Furthermore, this research is being conducted years after the shift to online distribution has been completed reducing any issues with the unfamiliar format. Current literature on the impact of the different methods of SET distribution on the mean evaluation scores remains unclear. This study will provide further evidence of the effect the transition has had on mean evaluation scores.

3. Description of the Data

The data collected examines student evaluations in principles and introductory courses offered by the Economics Department at the University of South Carolina. The state's flagship university, located in Columbia, South Carolina, offers 300 degree programs to more than 31,000 students (www.sc.edu). The Economics Department is located within the Darla Moore School of Business and offers the economics major through two schools, the School of Business and the College of Arts and Sciences. The data used in this study comes from three courses taught within the Economics Department: Principles of Microeconomics, Principles of Macroeconomics and Introduction to Economics. The principles courses are required for business and economics majors, along with economics minors, and are often taken when students are freshmen or sophomores. Students enrolled in the Introduction to Economics are business minors and those with a major requiring the course as a supplement to their program. Students enrolled in the introductory course have a wider variety of majors and level of interest, relative to the principles courses. The authors chose to include only lower level economics courses in order to ensure that the results are applicable not only to economics, but also other departments across the university offering introductory courses to both major and non-major students.

A trial run of the project occurred in spring 2014, including two introductory and one principles section. No changes were made to the procedures for the subsequent semesters. During the fall 2014 and spring 2015 semesters all instructors teaching non-honors sections of principles and introductory economics were contacted and asked to participate in the project. Instructors who agreed signed a consent form allowing researchers to administer an in-class SET during class time, along with the official university online SET. The consent form also allowed the researchers to receive the online evaluation results directly. In addition to the three sections collected in the spring 2014, fall 2014 data was collected from 13 of the 17 non-

honors sections offered and 15 of the 18 non-honors sections participated in spring 2015 for a total of 31 sections.

The online SET was administered using the same process that has been in place since fall 2010, when the transition from in-class to online distribution occurred. Students were contacted by email two weeks before final exam week to participate in the evaluation. The online SET remained open until the day finals began and students had the ability to complete the evaluation any day or time during this period. Those students who did not complete the evaluation received follow-up email reminders from the learning management system to complete the online evaluation.

The process for administering the in-class SET mirrored the traditional method used in the past by many institutions. Instructors were asked to provide a class period during the final two weeks of classes to have the evaluation administered; the same time frame online evaluations were open. The authors oversaw the distribution, completion, and collection of the evaluations, with additional graduate student support in large lecture sections. The in-class SET was distributed without prior announcement, only students attending class on the date and time chosen by the section instructor. The online and in-class SET questions were identical, only the distribution method changed. The in-class SET form can be found in Appendix I, with a similar format used for the online evaluation.

The SET is divided into five categories: (1) objectives of the course, (2) instructor performance, (3) learning and skill development, (4) course material, examinations, and assignments, and (5) other questions and written comments. The questions are rated on a scale from one to five with a one denoting the student strongly disagrees (very dissatisfied) and a five representing the student strongly agrees (very satisfied). The instructor receives the results after final grades have been submitted which includes the average rating and standard deviation for each question. A set of summary statistics, or summary indices, are also provided to allow the instructor quick feedback on each of the five categories and the evaluation as a whole. The summary indices are calculated by averaging across questions in a given category. The global index is the average rating of all questions on the SET.

A potential bias present in our study is that students may be less likely to complete the online evaluations if they have participated in the in-class version. This issue was addressed by informing students of a difference between the two SET. A script was read by the researcher administering the evaluations in-class which clearly informed students each version would be viewed differently. It is believed by the researchers that the use of the script and the gift cards encouraged students to take the in-class evaluation as serious as the official online version. The script read:

“We are handing out course evaluations for you to complete for ECON (Course Number), with (Instructor Name). Please use a Black Pen or a Number 2 Pencil. Once you have completed the evaluation bring the form forward. We will be raffling off Starbucks gift cards to those of you who have completed the evaluation. Please print your name and email address on the separate piece of paper to be entered into the drawing. Winners will be notified later this week. The instructor will receive aggregate results that contain no student names after final grades have been posted to my.sc.edu. It is important to note that these evaluations are not the same as the online

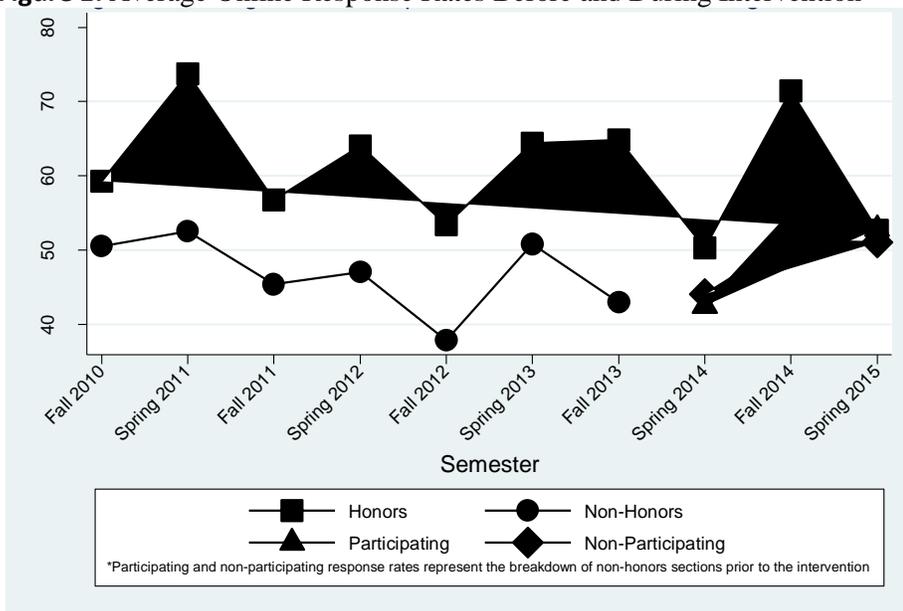
course and will be used in a different manner. You should still complete the online evaluations.”

Starbucks gift cards were offered to participating instructors to thank them for giving up class time to participate in the study. Gift cards were also awarded to a portion of students completing the in-class evaluation. Students were not aware of the reward before attending class and cards were not distributed until after the in-class evaluations had been completed. The authors wanted to be sure students took the evaluations seriously and completed all questions; the gift cards were used as a mechanism to encourage this.

The sections that chose not to participate, along with honors sections, are used to determine if the in-class evaluations discouraged students from participating in the online version. Online SET results from nonparticipating sections are not expected to be significantly different than the participating sections’ online SET results. Instructors of honors sections were not asked to participate in the project for a couple reasons. First, students enrolled in honors courses are already required to complete two online SET, adding a third in-class evaluation may create student fatigue. These sections were also not included in order to use them as a control group to investigate whether the intervention reduced participation rates for the online SET. While honors sections typically experience higher response rates, semester trends are similar across honors and non-honors sections.

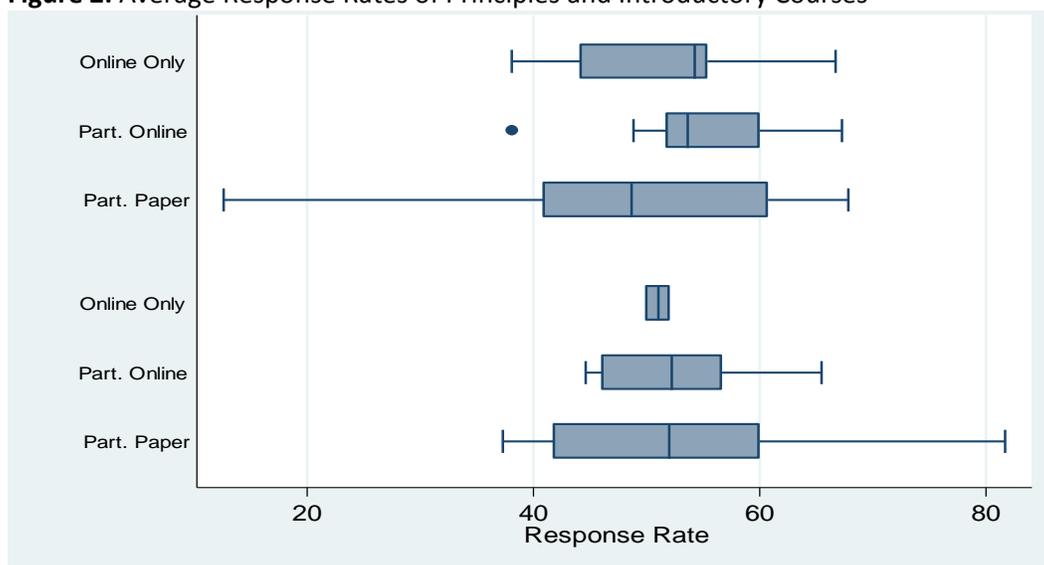
Figure 1 examines trends in the average online SET response rates for all principles and introductory sections from fall 2010, the semester the official transition to place, through the intervention semesters. If our intervention discouraged students from completing the online evaluations, the SET response rates for the participating sections would be much lower than that of the non-participating, non-honors sections and deviate from the trends present before the intervention period. The intervention includes the three semesters starting spring 2014 and ending spring 2015. For these semesters, the non-honors sections are divided based on whether they participated in the project. There is no significant difference in average response rates between the non-participating and participating sections during the intervention. During this period the honors sections maintain a significantly higher average response rate, though trends appear similar across groups, both during and prior to the intervention.

Figure 1: Average Online Response Rates Before and During Intervention



To further examine the impact the intervention had on response rates, the quartiles and ranges of both participating and non-participating, non-honors sections are examined. Figure 2 shows participation rates for the online SET for both non-participating and participating sections and in-class rates for participating sections. The figure further confirms that both non-participating and participating sections have similar medians. The fall 2014 online response rates show a similar range for non-participating and participating sections. Since there are only two sections that did not participate in the evaluations in spring 2015 there is little variance, but the medians are similar across groups. This figure further suggests the additional in-class evaluation did not discourage students from participating in the online evaluation.

Figure 2: Average Response Rates of Principles and Introductory Courses



The in-class median response rates are comparable to their online counterparts, but the range of response rates is much larger for the in-class SET. The span not only differs from the online response rates, but the range of in-class response rates vary across semesters. Fall 2014 experienced a wider range at the bottom of the distribution, while the reverse is true in spring 2015. This may be a semester effect; however the authors believe the result to be an instructor effect, likely caused by differences in attendance policies. While the majority of the instructors are overlapping across semesters, others only teach a principles course one semester an academic year. The differences in attendance policies are likely the cause of this variance.

4. Results

4.1. Results for the Full Sample

Using the aggregate SET results for both methods of distribution the researchers differenced the response rate, self-reported student characteristics, and average ratings for all specific questions and summary indices. The differenced equation is:

$$D_i = X_{c,i} - X_{o,i}$$

Where $X_{c,i}$ is the aggregate information obtained from the in-class SET in section i and $X_{o,i}$ is the aggregate information obtained from the online SET. A negative difference would suggest higher online rates or scores, while a positive would indicate in-class was greater. The differenced data was averaged across sections and tested against zero to measure whether there are significant differences across methods of distribution.

It is expected that online response rates will be significantly higher than in-class response rates. To participate in an in-class evaluation, the student must attend class on the particular day the evaluation is given. If the student arrives late or leaves early they may miss the survey. As the semester comes to a close attendance rates often fall, further decreasing participation. An online evaluation does not have these issues; all students enrolled in the course have access to the SET and are sent reminders to participate. The first row of Table 1 presents the results for the mean differenced response rate along with means for both the in-class and online individually. The expected results hold but are not significantly different. On average participation is two percentage points greater in online evaluations, relative to the in-class SET.

Table 1: Mean Descriptive Statistics for Course Evaluations

	(1) In-Class	(2) Online	(3) Difference
Participation Rate	51.03	53.07	-2.04
Expects an A	37.52	38.00	-0.48
Expects a B	44.72	41.84	2.88
Expects a C	16.52	17.67	-1.15
Expects to Fail	1.33	2.47	-1.14*
Required Course	95.95	95.76	0.19
Part of Major	60.64	62.45	-1.82
Freshman	15.55	15.53	0.01
Sophomore	55.97	54.93	1.04
Junior	19.97	21.19	-1.22
Senior	7.73	8.22	-0.50
Observations	31	31	31

* $p < .1$, ** $p < .05$, *** $p < .01$

Difference is taken by subtracting the online score from the in-class score

The method of distribution is expected to influence the composition of students completing the evaluation. It is predicted that those expecting to earn an A or B in the course will represent a larger percentage of in-class evaluations due to the positive correlation between attendance and course grade (Wienberg et al. 2009; McPherson et al. 2009). The results are as expected. Table 1 suggests that students expecting to earn a B represent a greater portion of in-class evaluations, while those expecting to fail make up a significantly greater portion of the online evaluations.

The majority of students enrolled in Principles of Microeconomics and Principles of Macroeconomics are business majors who need these economics courses as prerequisites for other courses, nearly all students enrolled in the course are freshman and sophomore standing. The number of juniors and seniors taking Introduction to Economics is slightly higher than those enrolled in the principles classes, but is still relatively small compared to the number of freshman and sophomores enrolled. The results show a greater proportion of in-class evaluations are completed by freshmen and sophomores although neither result is statistically significant.

The evaluation also inquires into whether the course is required and if it is part of the major. A larger portion of students who took the in-class evaluation said that the course is required while the reverse is true of those participating online, though neither variable is significantly different than zero.

Table 2 presents the differences between in-class and online evaluation ratings of a set of summary indices. The difference between the global index, a compilation of all questions included on the SET, is positive but not significantly different than zero. The lack of significance provides evidence that changing the method of distribution does not affect the average; overall evaluation scores instructors receive, following results from previous research (Avery et al.

2006; Dommeyer et al. 2004; Donovan et al. 2006). However, the global index is not the only score administrators consider when evaluating an instructor’s performance.

Table 2: Mean Descriptive Statistics for Global Index and Summary Indices

	(1) In-Class	(2) Online	(3) Difference
Global Index	3.94	3.94	0.00
Objectives of the Course	4.24	4.27	-0.03
Instructor Performance	4.14	4.09	0.05**
Learning and Skill Development	3.68	3.73	-0.05
Course Materials	3.91	3.95	-0.03
Other Questions	3.74	3.75	-0.01
Observations	31	31	31

* $p < .1$, ** $p < .05$, *** $p < .01$

Difference is taken by subtracting the online score from the in-class score.

Other summary indices in Table 2 suggest that students who took online evaluations gave slightly higher evaluations although none are significantly different from zero. The one summary index that is positive and statistically different from zero is the rating for Instructor Performance. Students participating in-class gave higher ratings than those who completed the evaluation online. This suggests that students who attended class were more satisfied by the instructor and viewed instruction positively. It also may indicate that students who stopped going to class were dissatisfied with the instructor. Other indicators viewed by instructors and administrators include specific questions about the instructor performance, learning and skill development, course materials, and other questions. Questions within each category may be emphasized more than the others. Often the ranking on the instructor’s knowledge of the material and whether the student would recommend the instructor are given more weight than whether the instructor explained the attendance policy and if assignments and assessments were returned in a timely manner. Table 3 presents the difference in a selection of specific questions within the summary indices deemed most important.

Table 3: Mean Descriptive Statistics for Specific Questions

	(1) In-Class	(2) Online	(3) Difference
Instructor Performance			
Instructor Availability	3.34	3.34	0.00
Instructor Well Prepared	4.31	4.22	0.09***
Instructor is Knowledgeable	4.39	4.31	0.07**
Effective Teaching Style	3.59	3.60	-0.01
Recommend this Instructor	3.71	3.70	0.00
Willingness to help	4.05	4.03	0.02
Learning and Skill Development			
Learned a great deal	3.72	3.74	-0.02
Useful later in Career	3.80	3.83	-0.02
Course Materials			
Textbook helpful	3.50	3.69	-0.19***
Consistent exams and assignments	4.04	4.03	0.01
Challenging exams and assignments	4.18	4.12	0.06
Other Question			
Interaction with Instructor	3.73	3.73	0.00
Observations	31	31	31

* $p < .1$, ** $p < .05$, *** $p < .01$

Difference is taken by subtracting the online score from the in-class score.

The first set of questions in Table 3 are taken from the instructor performance category. While the summary index of the instructor performance score is significantly higher for in-class evaluators, the ratings of specific questions are mixed but only significant in a limited number of cases. The instructor was consistently well prepared for the class and knowledgeable are the two cases in which the in-class evaluation scores are significantly higher than online.

Learning and skill development is the second category of questions presented. The results on whether the student learned a great deal and whether the topic will be useful in their career are negative, but not significantly different than zero. The other question about interactions with the instructor is also not significantly different than zero suggesting that there is little perceived difference between online and in-class SET.

The specific questions on course materials offer an interesting insight into some differences between the in-class and online evaluators. Those completing the online version are more likely to agree with the statement "The textbook and other course material contributed to my learning." It is suspected that students who have become discouraged with the course or instructor comprise a greater percentage of online evaluators. This result suggests that those students may rely more on the textbook and other materials rather than the lecture and instructor. The other two questions in this category concerning exams and assignments are not significantly different than zero.

4.2 Results by Course

The next set of tables, Tables 4 through 6, divide the evaluation results by course type principles and introductory economics. Principles of Microeconomics and Principles of Macroeconomics courses are analyzed together due to the similarities between the sections in instructional quality, assessments, and the type of students enrolled; it also helped keep the number of observations larger. Business and economics majors represent the majority of students enrolled in principles courses and most are freshmen or sophomores due to these classes being prerequisites, leading to a relatively homogeneous set of students with similar backgrounds regarding interests and courses. Students enrolled in these courses are required to earn a C or better in the course, but also must maintain a 3.0 GPA to remain in the program. Introduction to Economics is designed for business minors, a supplement to other majors, and used as a core Social Science course. For the minor students do need to earn a C or better, however for the general education requirement a minimum grade of D is accepted. Students fail the introductory course at a higher rate and often have difficulties understanding how the topic applies to their major. The differences in student characteristics across these two course types raise questions about whether the distribution method of the SET impacts these results differently.

Table 4: Mean Descriptive Statistics for Course Evaluations by Course Type

	Principles Sections			Introductory Sections		
	(1) In-Class	(2) Online	(3) Difference	(4) In-Class	(5) Online	(6) Difference
Participation Rate	52.9	51.3	1.61	49.27	54.73	-5.46
Expects an A	39.12	38.7	0.42	36.02	37.34	-1.33
Expects a B	42.37	42.35	0.02	46.93	41.36	5.57**
Expects a C	17.22	17.75	-0.53	15.86	17.6	-1.74
Expects to Fail	1.3	1.19	0.11	1.36	3.67	-2.31**
Required Course	97.01	97.47	-0.47	94.96	94.16	0.8
Part of Major	76.16	78.09	-1.93	46.08	47.79	-1.71
Freshman	20.37	21.59	-1.22	11.03	9.86	1.17
Sophomore	65.59	64.86	0.73	46.95	45.62	1.33
Junior	11.48	10.73	0.75	27.93	31	-3.07*
Senior	1.87	2.78	-0.91	13.22	13.32	-0.11
Observations	15	15	15	16	16	16

* $p < .1$, ** $p < .05$, *** $p < .01$

Difference is taken by subtracting the online score from the in-class score.

The first row of Table 4 presents the participation rates and differences in the online and in-class SET for the principles and introductory sections. There is no statistically significant difference between the two methods of SET for both the principles and introductory classes. However, there are significant differences in the attributes of students completing the SET for the introductory sections. A significantly higher portion of students complete the in-class evaluation expect to earn a B in the class, while a higher percentage of online evaluators expect

fail the class. There are also statistically more juniors in the introductory sections filling out the online SET. The lack of significant differences between participation rates and aggregate student attributes for the principles sections suggests that the method of distribution has a very small to no impact.

Table 5 presents the results for the global index and summary indices by course type. For both the principles and introductory sections the global index is not statistically different from zero. This again suggests that overall there is little difference between methods of delivery. For the principles sections the summary indices for learning and skill development and course materials is significantly higher for online respondents. For introductory sections, in-class respondents reported significantly higher scores for the objectives of the course and instructor performance summary indices.

Table 5: Mean Descriptive Statistics for Global Index and Summary Indices by Course Type

	Principles Sections			Introductory Sections		
	(1) In-Class	(2) Online	(3) Difference	(4) In-Class	(5) Online	(6) Difference
Global Index	3.94	3.97	-0.03	3.95	3.91	0.04
Objectives of the Course	4.24	4.39	-0.15	4.24	4.16	0.09*
Instructor Performance	4.14	4.12	0.02	4.14	4.07	0.08***
Learning and Skill Dev.	3.76	3.87	-0.11**	3.61	3.61	0
Course Materials	3.88	3.95	-0.07*	3.94	3.95	-0.01
Other Questions	3.68	3.73	-0.05	3.79	3.77	0.02
Observations	15	15	15	16	16	16

* $p < .1$, ** $p < .05$, *** $p < .01$

Difference is taken by subtracting the online score from the in-class score.

Table 6 presents the differences between specific questions for principles and introductory sections separately. These results further show differences by course type. For the principles courses those who took the online SET gave a significantly higher rating to the course being useful later in their career. There is also a significantly higher rating for the online SET about the helpfulness of the textbook. As before, students that may be unsatisfied with the instructor and do not come to class anymore may have to rely on the textbook more to get the information. Contrary to expectations, online SET respondents provided higher ratings on the interaction with the instructor. This may be due to these students relying on communication outside classroom.

Results for the introductory sections suggest that students who took the in-class SET rated the instructor as being well prepared and knowledgeable is statistically higher compared to the online SET results. Students who are in class to take the SET are more likely to be satisfied with the instructor and think they are prepared and knowledgeable; if this were not the case then they are likely to join the apathetic student who may not attend class regularly. As with the principles courses, the helpfulness of the textbook was higher for those who took the online SET. Again, this is likely because students who stopped attending class or who did not attend class regularly might have to rely on the textbook more. When dividing the data

based on course type the introductory sections appear to be driving the majority of results for the aggregate analysis. This may be due to the differences in motivation and level of interest across the course type.

Table 6: Mean Descriptive Statistics for Specific Questions by Course Type

	Principles Sections			Introductory Sections		
	(1) In-Class	(2) Online	(3) Diff.	(4) In-Class	(5) Online	(6) Diff.
Instructor Performance						
Instructor Availability	3.31	3.32	-0.01	3.37	3.36	0.01
Instructor Well Prepared	4.36	4.29	0.07	4.26	4.16	0.10***
Instructor is Knowledgeable	4.45	4.39	0.06	4.33	4.24	0.08***
Effective Teaching Style	3.7	3.73	-0.03	3.49	3.48	0.01
Recommend this Instructor	3.78	3.81	-0.03	3.64	3.61	0.04
Willingness to help	4.03	4.02	0.00	4.08	4.04	0.05
Learning and Skill Development						
Learned a great deal	3.79	3.86	-0.08	3.65	3.62	0.03
Useful later in Career	3.95	4.04	-0.09**	3.67	3.63	0.04
Course Materials						
Textbook helpful	3.41	3.66	-0.25***	3.59	3.72	-0.13*
Consistent exams & assignments	4	4.03	-0.03	4.07	4.03	0.05
Challenging exams & assignments	4.21	4.14	0.07	4.15	4.1	0.05
Other Question						
Interaction with Instructor	3.63	3.71	-0.07*	3.81	3.75	0.06
Observations	15	15	15	16	16	16

* $p < .1$, ** $p < .05$, *** $p < .01$

Difference is taken by subtracting the online score from the in-class score.

5. Conclusion

The transition from evaluations administered in-class to online sparked concern among instructors about lower response rates and lower SET mean scores. The data collected for this project is used to examine whether these worries are valid. In-class and online evaluations were administered in multiple sections of Principles of Microeconomics, Principles of Macroeconomics, and Introduction to Economics courses at the University of South Carolina. The in-class evaluations employed a process similar to that used before the transition online. The online evaluations are the official SET of the college and have been the standard since 2010. The primary goal of this research was to determine whether online distribution influenced mean SET scores and response rates in as negatively as some had perceived. The expectation was the method of distribution would lead to differences in access to the SET and, as a result, change the composition of students who participated.

Findings suggest no significant difference in participation rates between online and in-class distribution methods for sections as a whole. However, when the sample is divided by course type, a higher response rate for the online SET is observed for introductory sections. The difference in response rates alters the composition of students completing the evaluations. Students expecting a B represent a statistically significant greater portion of in-class evaluations, while those expecting to fail make up a larger portion of the online evaluators, specifically in introductory sections. Concerning the global index and summary indices, there are few significantly different results across methods of distribution. The one exception is the instructor performance category. Primarily driven by the introductory sections, in-class evaluators gave higher ratings to their instructors.

The results for specific questions on the SET are mixed. Inquiries into the instructor preparedness, knowledge, and the course having challenging exams and assignments are scored higher by in-class evaluators, while online ratings are higher for the usefulness of the textbook and other course materials. These differences are important to recognize as the results are often examined by instructors and administrators to measure effectiveness. Though not all results are significantly different across methods of distribution, the difference in ratings for these important indicators of teaching should be acknowledged.

The data in this study allows results from online and in-class SET of the same class to be estimated. However the data set used for this research does not directly include information on whether students responded to one, both, or neither of the evaluations. Nor does it have information on whether students give similar ratings under the different methods of distribution. In order to acquire candid responses the authors chose to keep the in-class evaluations anonymous, this does not allow for matching students who completed the in-class and online evaluations. Data containing this information could provide further insight into differences between the students and their responses to online and in-class SET.

It should also be noted that all student attributes analyzed in this paper are self-reported. There is a possibility that some students might not have correct expectations of their grades. However, since the SET is administered in the final weeks of class, the majority of students should have proper information to report an informed expected final grade. Since no student identifiers were included in the data, the actual grade earned is not available even after final grades are determined.

Another limitation is the lack of individual student data, only aggregate statistics are available. More information on specific student attributes and their rating of the course would be valuable towards understanding students' opinions of the course. However, this was not the goal of our study. In order to accurately measure the impact of the difference between the online and in-class evaluations on the mean SET results, it was important not to alter any component of the evaluation form. No changes were made to questions or the methods of distributing the SET to ensure results reflected only the difference across distributions methods.

The goal of this research is to provide insight into the effect the method of distribution has on response rates, mean instructor performance, and mean course ratings reported on SET. Results from this paper are not only important for the Economics Department at the University of South Carolina, but any department at any university or college that has transitioned to online distribution of the SET or plans to do so in the future. We find evidence that the response rates on the web-based evaluations will in time, if not initially, approach the response

rates from the in-class SET; this dispels one of the most common criticisms of administering SET online present in existing literature. Results also suggest that the method of distribution may have the greatest impact on introductory sections designed for non-majors.

Tenure and promotion decisions are partially based on student evaluations of teaching, since this is the case, it is important to be aware of any changes that will result from the transition. Traditional in-class evaluations require students to be present to participate in the evaluation, while the online distribution opens the evaluation to all students enrolled in the course. This appears to change the composition of students completing the SET, increasing the weight of dissatisfied and discouraged students no longer attending class.

The results presented in this research raise a greater question beyond that of the impact the method of distribution has on SET ratings. The question instructors and college administrators must first address is the goal of the SET. If the goal of the SET is to hear the opinion of all students enrolled in the course, the online method of distribution is best. However, some students may be disengaged from the course, not attending class or participating in any way. Do those viewing the SET results want to capture these students' opinions, while missing that of the apathetic student attending class, but lacking motivation to participate online? Determining the goal of the SET will guide the decision of which method of distribution is best for each college's situation.

References

Avery, R. J., W. K. Bryant, A. Mathios, H. Kang, and D. Bell. 2006. "Electronic Course Evaluations: Does an Online Delivery System Influence Student Evaluations?" *Journal of Economic Education* 31(1): 21-37

Becker, W.E., W. Bosshardt, M. Watts. 2012. "How Departments of Economics Evaluate Teaching," *Journal of Economic Education*. 43(3): 325-333

Carini, R., J. Hayek, G. Kuh, J. Kennedy and J. Ouimet. 2003. "College Student Responses to Web and Paper Surveys: Does Mode Matter?" *Research in Higher Education*. 44(1):1-19.

Dommeyer, C.J., P. Baum, R.W. Hanna and K.S. Chapman. 2004. "Gathering Faculty Teaching Evaluations by In-class and Online Surveys: Their Effects on Response Rates and Evaluations." *Assessment & Evaluation in Higher Education*, 29(5): 611-623

Donovan, J., C.E. Mader, and J. Shinsky. 2006. "Constructive Student Feedback: Online vs. Traditional Course Evaluations." *Journal of Interactive Online Learning*, 5(3): 283-296

Ewell, B. 2000. The United States Air Force Academy (AFA) desired to convert its mid- and end-of course evaluations from paper-pencil to computer administered methodology. Retrieved May 6, 2015 from <http://home.att.net/~bobewell/oleval.htm>

Fike, D.S., D.J. Doyle and R.J. Connelly. 2010. "Online vs. Paper Evaluations of Faculty: When Less Is Just As Good," *Journal of Effective Teaching*, 10(2):42-54

Hardy, N. 2003. "Online ratings: Fact or fiction." In D. L. Sorenson, and T. D. Johnson (Eds.), *New Directions for Teaching and Learning: Online Student Ratings of Instruction*, 96. San Francisco: Jossey-Bass.

Johnson, T. 2002. *Online Student Ratings: will Students respond?* Paper presented at: American Educational Research Association Annual Meeting, 2002.

Kherfi, S. 2011. "Whose Opinion Is It Anyway? Determinants of Participation in Student Evaluation of Teaching" *Journal of Economic Education*, 42(1): 19-30

Kronholm, E. A., Wisher, R.A., Curnow, C.K, and Poker, F. 1999. "The Transformation of a Distance Learning Enterprise to an Internet Base: From Advertising to Evaluation," paper presented at the Northern Arizona University NAU/Web99 Conference. Flagstaff, AZ

Layne, B. H., J.R. DeCristoforo, and D. McGinty. 1999. "Electronic versus traditional student ratings of instruction," *Research in Higher Education*. 40(2): 221–232

McPherson, M.A., R.T. Jewell and M. Kim. 2009. "What Determines Student Evaluation Scores? A Random Effects Analysis of Undergraduate Economics Classes" *Eastern Economic Journal*, 35:37–51

Morrison. K. 2013. "Online and Paper Evaluation of Courses: a Literature Review and Case Study". *Educational Research and Evaluation*, 19(7): 605-614

Morrison, R. 2011. "A comparison of online versus traditional student end-of-course critiques in resident courses" *Assessment & Evaluation in Higher Education*. 36 (6): 627-641

Opengart , R .and R.R. Mau. 2012. "Comparing Ratings: In-class (paper) vs. Out of Class (online) Student Evaluations" *Higher Education Studies*, 2(3):55-68

Stowell, J.R., W.E. Addison, J.L. Smith. 2012. "Comparison of Online and Classroom-based Student Evaluations of Instruction," *Assessment & Evaluation in Higher Education*, 37(4): 465-473