Do Students Really Learn from Online Experiments? – Evidence from Introductory Microeconomics Class[1,2]

S. Kumarappan[3]
Ohio State ATI
Wooster, OH

Abstract

Online experiments are frequently used to engage students and improve pedagogy in introductory microeconomics classrooms. This paper compared student scores on homework problem sets to evaluate whether the experiments helped improve student understanding of economic concepts. Two composite scores were created for each student: one based on the homework problem sets that involved an online experiment component and another for the problem sets that did not have any associated online experiments. The results showed an increase in student scores, ranging from 1% to 5%, when online experiments were conducted prior to the related homework. Two statistical tests – paired t-tests conducted on the mean score and Wilcoxon signed rank test on the median scores – showed that the increase in student scores are statistically significant. The variances in student scores were also found to be reduced with the use of online experiments. The descriptive answers given by students were able to recall the experimental setup and use it to explain related economic concepts.

Introduction

Agribusiness classrooms have long employed modern technologies such as newer presentation software, audio-visual tracks, online assessments and online experiments that simulate market and business settings and changed how students participate, engage and learn (Alston et al., 2003; Litzenberg, 1995; Litzenberg, 1982). New pedagogical models such as the Technological, Pedagogical, and Content Knowledge (TPACK) suggest that technology has become an essential component of classroom pedagogy and an integral part of subject content (Koehler and Mishra, 2009). Particularly, in undergraduate economics classrooms, technological resources are becoming more common than ever (Kennelly and Duffy, 2007; O'Dea and Ring, 2008).

The basic economic concepts can be taught with or without the use of technology such as experiments (Becker and Watts, 1995; Carter and Emerson, 2012; Emerson, 2014; Joseph, 1970; Wells, 1991). Nguyen and Trimarchi (2010) reported student performance improved slightly but significantly, showing a marginal (two percentage point) increase in student grades with the use of homework software packages (such as Aplia or MyEconLab). The software packages help organize and provide easy access to course content; can it also help improve students' economic knowledge? The existing studies find mixed evidence with the use of technology in teaching economics through synchronous online experiments (O'Dea and Ring, 2008; Lee et al., 2010; Perez-Sebastian, 2010). Technology has the advantages of increasing enthusiastic participation of students and highlight the nuances of economic concepts through quick implementation of multiple rounds with slight variations and immediate summarization of results (Ball and Eckel, 2004; Janssen et al., 2014; Palan, 2014; Shor, 2003). But, the question of whether students become savvy in the economic content due to online experiments is yet to be investigated. This study tries to find evidence for any discernible improvements in student performance (grades) when technological tools (such as online experiments) are used to teach key introductory economic concepts.

The students in introductory microeconomics class need to know the impact of demand and supply forces and how they determine the outcome in these issues: (i) laissez faire market experiments show what the market equilibrium would look like (quantity and price) under free market conditions, (ii) government intervention measures that can have negatively implications such as surplus or shortage and deadweight loss, (iii) government intervention that is necessary to manage of common pool resources, (iv) how monopoly profits and deadweight loss can be controlled by the government and (v) how information asymmetry problems can be addressed with rules and requirements mandated by the government or markets. Each of these issues builds upon each other to provide a broader economic

[1]This study was deemed exempt by the Ohio State University Institutional Review Board.
[2]The homework problem sets show a slightly higher score due to the availability of three attempts for each question.
[3]Division of Arts, Sciences, and Business; Phone:330-287-1261; Email: kumarappan.1@osu.edu

understanding for the students. These concepts can be taught with the use of online experiments, as explained below.

While using online experiments, the class is divided into two sections – buyers and sellers. The buyer group will be provided with a buyers' value signifying their willingness to pay; the seller group will be provided with sellers' costs signifying their willingness to accept. The buyers and sellers are given with decide how much to sell for or bid on the product. The technological component makes it convenient to implement various rounds with slight variations. One of the primary benefits of using online experiments is that it can provide students with additional time to think and reflect upon the economic concepts and understand the economic implications better. The online experiments used for this study comes from multiple sources such as Aplia, MobLab, MyEconLab, GameTheory.net and VEconLab.

This study hypothesizes that when online experiments are used, they can reinforce the economic concepts in students and help them score better in their assessments. The hypothesis can be established by comparing two scores: (i) the students' performance in homework problem sets that are completed after online experiments and (ii) the students' performance in homework problem sets that did not include any prior online experiment component (Bostian and Holt, 2013; Holt, 2009; Nguyen and Trimarchi, 2010; Shor, 2001). The results discussed below show that there is a marginal improvement in students' scores, which can be considered as a proxy of gain in students' economic knowledge and understanding. The improvement in student scores is statistically significant. The following sections provide an overview and nature of data, methods and the results and implications.

Data

The data for this analysis were derived from an introductory microeconomics course (*titled Principles of Food and Resource Economics*) in the Ohio State University ATI, Wooster, OH over four semesters – spring 2013-fall 2014. The delivery format of all four course offerings remained the same: lectures supplemented with audio-visual aids and experiments. The assessments included weekly homework problems and exams. Each offering of the course contained 12 problem sets. On an average, students completed four homework problem sets after participating in a related online experiment that reinforced learning; other eight problem sets did not have any underlying experiment. The class sizes ranged from 59-97 (table 1); the students' mean and median scores for the entire course grade (including exams) ranged from 77 to 82%. The mean and median of homework problem set scores ranged from 82-88% and 93-97% respectively (see footnote 2).

Table 1: Summary statistics for the introductory microeconomics course

		Overall course grade		Problem set grade	
Term	Number of students	Mean (%)	Median (%)	Mean (%)	Median (%)
Fall 2014	97	80	82	82	93
Spring 2014	83	77	81	85	97
Fall 2013	59	78	81	88	97
Spring 2013	69	78	81	83	97

The online experiments, implemented primarily through Aplia software, included basic economic concepts such as finding market equilibrium, evaluating the role of the government (price controls, impact of taxes), managing common resources (tragedy of the commons) and making decisions in the presence of information asymmetry. The classroom attendance rate, which is an indicator of students' effort level, was similar irrespective of the presence or absence of experiments: the attendance rate ranged from 86-93% when experiments were conducted and 85-94% for general lectures. Given that each student's effort level remained the same, any difference in student problem set scores would reflect the knowledge gained from the online experiment and how that knowledge got translated into better grades in the homework problem sets.

The composite data created for each student is paired in nature. That is, for each student, the factors such as effort, interest level and preliminary knowledge remain the same. The two composite scores for each student differed primarily in that whether the problem set was completed with or without prior online experiment component. Hence, comparing the changes in problem set score for each student individually is a reliable way to evaluate if the students gained knowledge from participating in online experiments. Such a paired nature of data also eliminates the need to control for other variables such as students' knowledge level (measured by GPA, SAT scores), status or year in the college (freshmen, sophomore), gender, race, age and prior economics courses completed, effort level (attendance) and other factors as reported in Carter and Emerson (2012). These are some of the factors that can possibly explain student scores – but the paired data eliminates the need to control these factors. Hence, analyzing the composite scores for each student, across four semesters, presents a reliable way to evaluate the pedagogical effectiveness of online experiments.

Methods

For each student, two composite homework scores were created: y_i denotes the composite score for student i with underlying online experiments and w_i denotes the composite score for the problem sets that did not have any underlying experiments. The difference in scores for each student be represented by $x_i = y_i - w_i$. The data for each student y_i and w_i would be independently distributed; that is, the composite scores derived for a student depend only the effort and knowledge level of that particular student and independent of other students. If students did not benefit from experiments, then the expected value of x_i, denoted as would be zero. That serves as the null hypothesis for testing, = 0. The alternative hypothesis is ≠ 0; it allows for both > 0 where the online experiments have a positive impact on student knowledge and < 0 where the online experiments have a negative impact on student knowledge.

The descriptive statistics for the mean and median scores in table 1 displayed skewness in student score distribution. To correct for the skewness, the data was given a monotonic log-transformation. The sample size for each class ranged from 59 to 97 students, sufficient enough for the variable xi to be t-distributed. To evaluate whether the mean value of variable x_i, , was zero or not, t-statistic was computed as t = (- μ)/ SD where was the average difference in the scores with and without online experiments; μ was the mean difference assumed to be zero under the null hypothesis and SD was the standard deviation of the differences in student scores. The computed value of t-statistic could be compared against the t-critical value for a two-tailed distribution to allow for both possibilities > 0 and < 0.

A non-parametric test known as Wilcoxon signed rank test for paired samples was also conducted. The key idea behind this test was to test whether more than half of the class improved their homework problem set scores with the use of online experiments. Hence, it was a test on the median score of the students' problem sets. Both the required criteria for the use of Wilcoxon signed rank test were satisfied by this data set: (i) x_i values were independent of each other and (ii) y_i and w_i were interval data (to enable ranking of student scores based on the difference of y_i and w_i). The steps to conduct the Wilcoxon signed rank test were as following: First the differences in the student scores $x_i = y_i - w_i$ was calculated. Second, a rank score was assigned to each student in the class based on the absolute value for x_i. Third, two groups of students were created based on the sign of x_i (that is x_i >0 and x_i <0). Finally, the rank values were summed up for both groups. To test for statistical significance, the lower value of the two sums (called as the z-score) was compared against the critical value available from the table for Wilcoxon signed rank tests. More details and a ready-to-implement spreadsheet tool were available with Zaiontz (2014).

Results and Implications

Table 2 shows that the mean score from problem sets increased in all four semesters. The online experiments help increase the homework scores by a nominal amount of 1% to 5%. Though marginal, the improvement in raw scores is found to be statistically significant for fall 2013 and spring 2013 semesters; the paired t-test statistic values are 2.798 and 5.234 which are significant at 1% level. The statistical significance can be taken as evidence that underlying online experiments help improve students' knowledge of economic concepts. The slight increase is reflective of previous studies that involved online homework software (especially, Nguyen and Trimarchi, 2010). The increases in students' mean scores in spring 2014 and fall 2014 semesters have rather been marginal and non-significant. One possible reason could be the larger class sizes (table 1) compared to other classes. Even though the paired t-test does not show statistical significance, the Wilcoxon signed rank test shows that the improvement in student scores is statistically significant during all semesters as given below.

The Wilcoxon signed rank test results are presented in table 3. The computed z-scores range from 1.7409 in fall 2014 to 5.452 in spring 2013. The results are significant at 1% level during spring 2013 and fall 2013; and at 5% and 10% levels during spring 2014 and fall 2014 semesters. The Wilcoxon signed rank test results show that more than half (median score) of the students have been able to improve their homework scores upon completing a related online experiment.

Coefficient of variance: The positive effect of online experiments established through the statistical significance is a direct result of higher mean composite score and lower coefficient of variation. Hence, the problem sets with associated online experiments would portray higher mean composite scores or lower standard deviation or both. See figure 1. A cursory look at the chart also reveals that the problem sets with underlying experiments (denoted by triangles) and without experiments (denoted by dots) display a higher mean score or less variance or both. The vertical line signifies the average (mean) score for all problem sets during that semester; the horizontal line is the average measure of standard deviation during that semester. Any score that is below the horizontal line or to the right of vertical line – falling in the bottom right corner – can be considered ideal; or at the very least, the scores should fall to the right of the vertical line. The problem sets with ideal outcomes were three out of five in spring 2014, two out of three in fall 2013 and all the three out of three problem sets associated with experiments in spring 2013. During fall 2014, only one of the four problem sets associated with experiments (denoted by triangles in the figure) reflect the ideal outcome of higher mean or lower coefficient of variation or both.

Table 2: Paired t-test results for the difference in mean composite scores with and without experiments

Term	Mean composite score for problem sets without underlying experiments ()	Mean composite score for problem sets with underlying experiments ()	Mean value^ of	t-statistic
Fall 2014	81%	82%	1% (0.013)	0.894
Spring 2014	84%	85%	1% (0.014)	0.625
Fall 2013	86%	90%	4% (0.015)	2.708 ***
Spring 2013	82%	87%	5% (0.010)	5.234 ***

^ positive values of show that problem set scores increased with experiments; standard error in the parenthesis
*** significant at 1% level ; ** significant at 5% level; * significant at 10% level

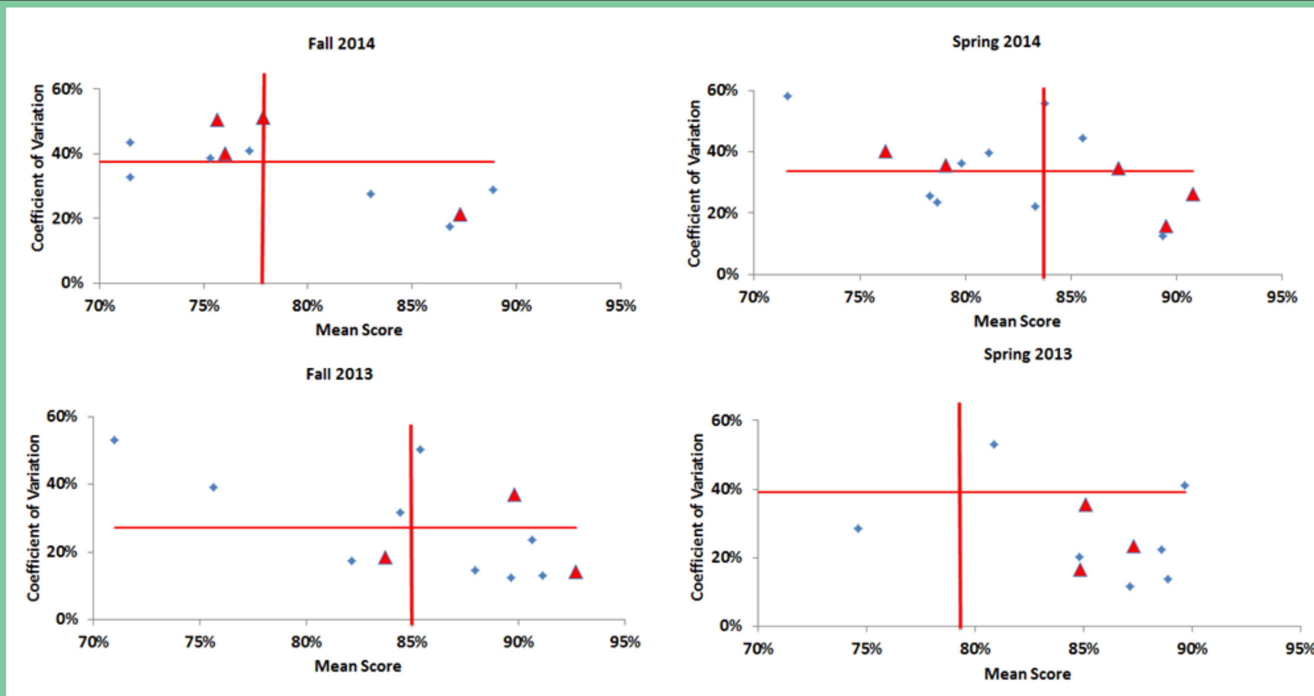
Table 3: Wilcoxon signed rank test results for the difference in median composite scores with and without experiments

Term	Sum of Ranks#	T-critical value for rank	z-score
Fall 2014	T- = 1811; T+ = 2749	1751	1.741 *
Spring 2014	T- = 1189; T+ = 2214	1277	2.369 **
Fall 2013	T- = 414; T+ = 1297	602	3.418 ***
Spring 2013	T- = 252; T+ = 1959	798	5.452 ***

The lower value of the sum is compared with the critical value; if the lower value of the sum is less than the critical value, then the null hypothesis is rejected.

T- and T+, respectively, refers to the sum of ranks of students whose problem set scores were lower and higher when underlying experiments were employed.

Figure 1: Distribution of individual homework scores^



^ The triangles and dots represent problem sets with and without underlying experiments, respectively. The vertical line is the mean score of all problem sets irrespective of the presence of absence of experiments; the horizontal line is the average value of coefficient of variation of all problem sets.

Qualitative assessment: The differences across different student batches are apparent: hence the student cohort (and possibly class sizes) could affect how well students benefit from online experiments. The students' descriptive answers for essay questions suggest that students are able to recall the setup of online experiments and recognize its economic significance. According to Carter and Emerson (2012), online experiments are as effective as manual experiments; according to this study, online experiments are effective in improving student scores. Overall, the results and observations presented here provide further evidence that online experiments can be a useful pedagogical tool in economics classrooms. There is a caveat to these results. The students completed the homework problem sets within a few days after the online experiment; hence, the results presented here could be more indicative of short term knowledge retention. The long run benefits can be established by including suitable exam questions in the final exam, after a substantial gap of participating in the online experiments.

Conclusions

This paper evaluated whether online experiments helped improve students' economic knowledge. The data from homework problem sets in four introductory microeconomics courses during spring 2013 through fall 2014 were used. Each student's homework was used to derive two composite scores – one based on the homework that included underlying online experiments and another composite score without any online experiment component. The summary statistics showed that mean scores were slightly higher by 1% to 5% when online experiments were employed as part of classroom teaching. The paired t-tests conducted on the mean scores and Wilcoxon signed rank test on the median scores showed statistical evidence (at varying levels) for improvement in student performance. In addition to higher mean score, the variation in student response was lower. The students' descriptive answers in exams showed better understanding of economic concepts when online experiments were employed for classroom instruction.

Literature Cited:

Alston, A.J., W.W. Miller and D.L. Williams. 2003. The future role of instructional technology in agricultural education in North Carolina and Virginia. Journal of Agricultural Education 44(2): 38-49.

Ball, S.B. and C.C. Eckel. 2004. Using technology to facilitate active learning in economics through experiments. Social Science Computer Review 22(4): 469-478.

Becker, W.E. and M. Watts. 1995. Teaching tools: Teaching methods in undergraduate economics. Economic Inquiry 33(4): 692-700.

Bostian, A.A. and C.A. Holt. 2013. Veconlab classroom clicker games: The wisdom of crowds and the winner's curse. The Journal of Economic Education 44(3): 217-229.

Carter, L.K. and T.L. Emerson. 2012. In-class vs. online experiments: Is there a difference? The Journal of Economic Education 43(1): 4-18.

Emerson, T.L. 2014. Anyone? Anyone? A guide to submissions on classroom experiments. The Journal of Economic Education 45(2): 174-179.

Holt, C. 2009. University of Virginia Veconlab. Available at: http://veconlab.econ.virginia.edu/admin.htm.

Janssen, M.A., A. Lee, T. Waring and D. Galafassi. 2014. Experimental platforms for behavioral experiments on social-ecological systems. Ecology and Society 19(4): 20.

Joseph, M.L. 1970. Game and simulation experiments. Journal of Economic Education 1(2): 91-96.

Kennelly, B. and D. Duffy. 2007. Using Aplia software to teach principles of economics. Development in Economics of Education Annual Conference. Cambridge, UK.

Koehler, M. and P. Mishra. 2009. What is Technological Pedagogical Content Knowledge (TPACK)? Contemporary Issues in Technology and Teacher Education 9(1): 60-70.

Lee, W., R.H. Courtney and S.J. Balassi. 2010. Do online homework tools improve student results in principles of microeconomics courses? American Economic Review (2010): 283-286.

Litzenberg, K.K. 1995. Agribusiness industry expectations of computer skills of agricultural economics and agribusiness students. Journal of Agriculture and Applied Economics 27(01): 104-111.

Litzenberg, K.K. 1982. Computer use in the agricultural economics classroom. American Journal of Agricultural Economics 64(5): 970-977.

Nguyen, T.T. and A. Trimarchi. 2010. Active learning in introductory economics: Do MyEconLab and Aplia make any difference? International Journal for the Scholarship of Teaching and Learning 4(1): 10.

O'Dea, W.P. and D. Ring, 2008. The impact of Aplia on student performance in intermediate microeconomics theory. American Economic Association Annual Meeting. New Orleans, LA.

Palan, S. 2014. A software for asset market experiments. No. 2014-01. Faculty of Social and Economic Sciences, Karl-Franzens-University Graz.

Perez-Sebastian, F. 2010. The use of experiments in macroeconomics courses. EDULEARN10 Proceedings: 5640-5643.

Shor, M. 2001. GameTheory. net: Repeated Prisoners' Dilemma Applet. http://www.gametheory.net/web/pdilemma/.

Shor, M. 2003. Game theory. net. The Journal of Economic Education. 34(4): 388-388.

Wells, D.A. 1991. Laboratory experiments for undergraduate instruction in economics. Journal of Economic Education. 293-300.

Zaiontz, C. 2014. Real statistics using Excel. http://www.real-statistics.com/free-download/real-statistics-resource-pack/.