Classroom Experiments: Is more more?

By Tisha L.N. Emerson and Linda English*

* Emerson: Department of Economics, Baylor University, One Bear Place #98003, Waco, TX 76798-8003 (e-mail: Tisha_Nakao@baylor.edu).

Over the last 20 years, a considerable literature on the efficacy of classroom economics experiments in promoting student achievement has developed. Researchers find that student achievement is generally enhanced (or at least unharmed) by exposure to a pedagogy utilizing classroom experiments. While these findings have encouraged the adoption of classroom experiments, questions remain - perhaps most notably - regarding the appropriate number of experiments.

In the current study we attempt to determine whether more intensive use of classroom experiments is associated with greater student achievement. Our data, collected between 2002 and 2013, contains variation in the number of experiments administered in a principles of microeconomics course. This variation allows us to estimate the impact on student achievement associated with exposure to different numbers of experiments.

I. Literature Overview

Published studies on the efficacy of classroom experiments range in their use from four to eleven experiments. At the lower end, Cardell et al. (1996) administer four experiments and find no significant effect on students' improvement on the Test of Understanding in College Economics (TUCE) over the course of the semester as compared to students in the no experiment control group. Yandell (1999) reports the findings of a study where students are exposed to two or six experiments. He finds no statistically significant difference on final exam performance between the two exposures.

Two studies administer a total of seven experiments to their treatment group. Dickie (2006) finds that, relative to the no experiment group, students exposed to experiments improved their TUCE scores over the semester, but only if their course grade was not dependent upon their experiment performance. Ball et al. (2006) also administered seven experiments and find that students in the treatment group earned significantly higher scores on the final exam, as compared to the no experiment control group.

Durham et al. (2007) administer a total of eight experiments to their treatment group. Student performance on topic-specific exam questions served as the measure of student performance. Durham et al. find mixed results with the treatment group outperforming, underperforming, and performing comparably to the control (no experiment) group depending on the topic.

At the upper end of experiment treatment intensity, Emerson and Taylor (2004) employ eleven experiments. While they find no significant difference between the control and treatment groups on 20 common final exam questions, they do find that the treatment group significantly outperforms the control on TUCE score improvement.

II. Data and Empirical Methodology

Students in our study enrolled in one of 28 sections of principles of microeconomics at Baylor University between the spring of 2002 and the fall of 2013. When registering for the course, students were not aware of the pedagogical approach that would be adopted in their particular section of the class nor were they aware that their class might be included in a study. The treatment group is comprised of fifteen sections that made extensive use of classroom experiments. Between six and eleven experiments were administered in these sections during the semester long course. The remaining thirteen sections made no use of experiments and, thus, constitute the control group. A traditional, lecture-oriented, chalk-and-talk pedagogical approach was adopted in all control sections where additional lecture was provided in place of experiments. Aside from the difference in the basic pedagogical approach (experiments vs. chalk-and-talk), care was taken to maintain consistency across the study sections. All 28 sections were taught by one of three instructors where each instructor taught both control and treatment sections.[1] Similar textbooks, homework assignments, quizzes, two to three midterm exams, and a comprehensive final exam were assigned in all sections. Exams were composed of a mixture of multiple choice, short answer and problem questions. Class enrollment for the study sections ranged from 14 to 170, with the vast majority in the 25-40 student range. Finally, all study sections were taught at similar times, starting between 9:00am and 12:30pm.

A. Model of Student Learning

We employ an educational production function approach that is standard in the literature (see, e.g., Siegfried and Fels, 1979) to motivate our empirical analysis. In this

[1] One instructor taught twelve sections where half were part of the treatment group. Another instructor taught four sections with one in the treatment group. The final instructor taught twelve sections with eight in the treatment group.

approach, the following reduced-form model is specified:

(1) Student learning = *f* (aptitude; educational background; other student-specific characteristics; educational environment or teaching methodology; professor and year controls)

Our student learning (achievement) measure is a student's final course score. Inputs to the educational production function include student aptitude (e.g., measured by students' combined math and verbal SAT score), educational background (e.g., whether a student has taken high school economics or is retaking the microeconomic course), class attendance, time constraints (current semester course load), other student-specific characteristics (e.g., age, gender and ethnicity), and the variable of interest: the number of experiments in which students participate over the course of the semester. Since the specifications control for student attendance, the number of experiments in which a student participated differs due to variation in the number of experiments administered in the course and the extent to which a student's absences occurred on days in which experiments were administered.

B. Student Achievement and Characteristics

The final course score for each student in the study depended upon their performance on homework assignments, quizzes, midterms, and a comprehensive final exam. Final course scores were calculated as the percentage of the total possible points that students earned on all assignments (homework, quizzes, and exams) over the semester.[2] Summary statistics for student achievement measures are presented in Table 1. There was no significant difference in mean course scores between students in the treatment and control groups.

TABLE 1— DESCRIPTIVE STATISTICS

	Full Sample	Control Group (No Experiments)	Treatment Group
Course Score	74.14	72.97	75.05
	(16.68)	(16.70)	(16.62)
Male	0.56	0.63	0.52†
	---	---	---
Age	19.75	19.76	19.75
	(0.97)	(1.03)	(0.91)
Nonwhite	0.28	0.28	0.28
	---	---	---
SAT Total	1171.41	1175.25	1168.40
	(125.50)	(117.64)	(131.38)
Number of Absences	2.59	2.74	2.48
	(3.14)	(3.18)	(3.11)
Repeating Course	0.09	0.08	0.11†
	---	---	---
High School Economics course	0.83	0.84	0.82
	---	---	---
Business Major	0.91	0.91	0.91
	---	---	---
Current Semester Hours	15.20	15.09	15.29
	(2.33)	(1.67)	493
N	880	387	15.29

†Treatment group's means are statistically different from the control group at the 5% (two-tailed) significance level or better.

[2] While the professors in the study had some differences in the number of homework assignments, quizzes and exams, the weighting across the assignments was similar.

To accurately estimate the effect of experiment participation it is important to control for individual student-level characteristics including gender, age, aptitude, effort, and socio-economic status. A portion of this data (gender, ethnicity, age, high school economics background) is collected through pre-course student surveys. Other measures (student SAT/ACT scores) are collected directly from student records for greater accuracy. Descriptive statistics for student characteristics are presented by treatment group status in Table 1. Slightly less than half (44%) of the sample was in the control group and thus exposed to no experiments. The remainder, the treatment group, participated in between one and eleven experiments with a mean of 5.5. The control group was more heavily male (63% as compared to 52% male in the treatment group) and slightly less likely to be repeating the course (8% vs. 11%), but otherwise was similar to the treatment group. The average student in our sample was between 19 and 20 years of age, majoring in business (91%), and enrolled in 15 credit hours at the time of the study. Slightly more than a quarter of the sample was an ethnic minority (28%). The average SAT score was 1171 and over three-quarters of the sample (83%) had taken an economics course in high school. Students averaged 2.5 absences over the course of the semester.

III. Experiment Participation and Overall Student Achievement

We estimate the effect of participation in experiments on our measures of student achievement for our sample of 880 students. Table 2 presents estimation results. We find a statistically significant positive relationship between the number of experiments in which a student participates and their course score. The impact is diminishing, however, as the number of experiments increases.

Participation in classroom experiments differentially impacts some groups. While older students outperform younger ones, the youthful disadvantage is partially offset by participation in experiments. Similarly, our findings suggest that an achievement gap exists between whites and ethnic minorities, but that experiments serve to help bridge this gap as well. In both cases, experiments may be providing market experiences that these groups (younger and minority students) may not otherwise have had and thus greater understanding of economic concepts that can't be gleaned simply from class discussion. As such, the use of classroom experiments may help reach these otherwise disadvantaged groups. However, experiments do not appear to differentially impact student performance by gender, aptitude, or attendance.

TABLE 2— COURSE SCORE (PERCENTAGE OF TOTAL AVAILABLE POINTS EARNED BY STUDENT)

Variable	Course Score
Number of Experiments	9.948***
	(3.561)
Number of Experiments^2	-0.154*
	(0.0767)
Male	-0.670
	(1.356)
Male x Number of Experiments	-0.171
	(0.260)
Age	3.046***
	(0.519)
Age x Number of Experiments	-0.444***
	(0.117)
Nonwhite	-3.200**
	(1.410)
Nonwhite x Number of Experiments	0.663**
	(0.299)
SAT Total	0.0506***
	(0.00568)
SAT Total x Number of Experiments	-0.000186
	(0.00117)
Absences	-3.521***
	(0.345)
Absences x Number of Experiments	0.201
	(0.136)
Repeating Course	2.116
	(1.789)
High School Economics	2.184**
	(0.885)
Business Major	-2.815
	(1.775)
Current Semester Hours	-0.360
	(0.256)
Observations	880
Adjusted R^2	0.487

Notes: Standard errors reported in parentheses are clustered at the class section level. Additional controls also included for professor, year, but omitted from the table.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

IV. Conclusion

Participation in classroom experiments has a positive, but diminishing, marginal benefit on students' final course scores. So, to an extent, "more is more" in the use of classroom experiments. Further, we find that classroom experiments can bridge some achievement gaps (between older and younger students and between whites and minorities).

REFERENCES

Ball, Sheryl, Catherine Eckel, and Christian Rojas. 2006. "Technology Improves Learning in Large Principles of Economics Classes: Using our WITS." American Economic Review 96(2): 442-446.

Becker, William, and John Powers. 2001. "Student Performance, Attrition, and Class Size Given Missing Student Data." Economics of Education Review 20: 377-88.

Bergstrom, Theodore, and John Miller. 2000. Experiments with Economic Principles: Microeconomics, McGraw-Hill Higher Education: USA.

Cardell, N. Scott, Rodney Fort, Wayne Joerding, Fred Inaba, David Lamoreaux, Robert Rosenman, Ernst Stromsdorfer, and Robin Bartlett. 2006. "Laboratory-Based Experimental and Demonstration Initiatives in Teaching Undergraduate Economics." American Economic Review Papers and Proceedings 86(2): 454-9.

Dickie, Mark. 2006. "Experimenting on Classroom Experiments: Do They Increase Learning in Introductory Microeconomics?" Journal of Economic Education 37: 267-88.

Durham, Yvonne, Thomas McKinnon, and Craig Schulman. 2007. "Classroom Experiments: Not Just Fun and Games" Economic Inquiry 45(1): 162-78.

Emerson, Tisha, and Beck Taylor. 2004. "Comparing Student Achievement Across Experimental and Lecture-Oriented Sections of a Principles of Microeconomics Course" Southern Economic Journal, 70(3): 672-693.

Maxwell, Nan, and Jane Lopus. 1994. "The Lake Wobegon Effect in Student Self-Reported Data." American Economic Review Papers and Proceedings, 84(2): 201-5.

Yandell, Dirk. 1999. "Effects of integration and classroom experiments on student learning and satisfaction" Proceedings: Economics and the Classroom Conference, Pocatello, ID: Idaho State University and Prentice-Hall Publishing Company: 4-11.