Very Preliminary Draft
Please do not cite without author's permission.

In-Class vs. Online Experiments: Is there a Difference?

Tisha L. N. Emerson*

December 2008

Abstract

The pedagogical approach of employing experiments in economics courses continues to increase in popularity. While original experiments were hand-run in-class, computerized, online experiments are now also available. Using a quasi-experimental approach, we investigate whether any difference in student achievement (as measured by the Test of Understanding in College Economics and final course scores) or other outcome measures exists between students exposed to experiments in-class and online. Students are "randomly" placed in one of three microeconomics principles classes. All classes have the same instructor and follow an identical syllabus. The only difference between the classes is the medium through which the experiments are administered. We find that those students who participate in hand-run, in-class experiments have significantly higher achievement and more favorable views of the experiments than those students who participated in the computerized, online experiments.

JEL code: A22
Keywords: economic education, experiments

---

* Tisha Emerson is an Associate Professor in the Department of Economics at Baylor University (One Bear Place, # 98003, Waco, TX 76798-8003; Phone: 254.710.4180; Fax: 254.710.6142; Tisha_Nakao@baylor.edu). The author gratefully acknowledges financial support from Baylor University. The patience and hard work during the data collection process of Susan Armstrong and the students who participated in our study are acknowledged with great gratitude. Finally, the paper has benefited greatly from comments from Steve Conroy. Any errors are the author's sole responsibility.

In-Class vs. Online Experiments: Is there a Difference?

Abstract: The pedagogical approach of employing experiments in economics courses continues to increase in popularity. While original experiments were hand-run in-class, computerized, online experiments are now also available. Using a quasi-experimental approach, we investigate whether any difference in student achievement (as measured by the Test of Understanding in College Economics and final course scores) or other outcome measures exists between students exposed to experiments in-class and online. Students are "randomly" placed in one of three microeconomics principles classes. All classes have the same instructor and follow an identical syllabus. The only difference between the classes is the medium through which the experiments are administered. We find that those students who participate in hand-run, in-class experiments have significantly higher achievement and more favorable views of the experiments than those students who participated in the computerized, online experiments. (*JEL* A22)

I. Introduction

In recent years, active learning techniques have been adopted at an increasing rate to either replace or augment the standard lecture-oriented, "chalk-and-talk" approach to teaching principles of microeconomics. One particularly popular active learning approach is the classroom experiment. Many of the experiments simulate markets allowing students to participate in transactions and experience the very phenomenon being discussed in their classes. Helping spur the adoption of the experimental pedagogy are recent studies that provide evidence that the experimental pedagogy results in higher student achievement (Emerson and Taylor 2004, Emerson and Taylor 2007, Dickie 2006) as compared with the traditional "chalk-and-talk" pedagogy.[1]

The Emerson and Taylor (2004, 2007) and Dickie (2006) studies employ in-class, hand-run experiments like those from the Bergstrom and Miller (2000) textbook, *Experiments with Economic Principles: Microeconomics*. While the hand-run experiments are well-developed and fairly simple to implement, they do have some shortcomings that make their adoption more difficult, if not impossible, in some classroom environments. For example, the hand-run, pencil and paper experiments work best for class sizes of 20-40 students. Beyond 50 to 60 students, the

[1] For further discussion of the literature on experiments see Emerson and Taylor (2004), Dickie (2006), and Durham et al. (2007).

manual administration of these experiments becomes quite challenging. Thus, instructors with large class sizes will face considerable difficulty in administering these hand-run experiments in their classes. Even with "manageable" class sizes, instructors may find record keeping and entering of data costly without an assistant.

Over time, computerized versions of many experiments were also developed and are now as widely available as the original, hand-run versions. The development of computerized (online or on a local server) experiments has made the experimental pedagogy accessible to a wider range of instructors as these computerized experiments can easily manage very large classes and automatically record data. Computerized experiments are available through a number of sources (both for a fee and free) including Aplia, VeconLab, and EconPort. As a result, one of the many variables an instructor must select when they decide to adopt the experimental pedagogy is the medium through which they will administer their experiments.

"Devotees" of the original hand-run experiments, however, question whether these computerized experiments afford students an experience equivalent to the hand-run experiments. Two important factors differentiate in-class and out-side of class (online) experiments. First, in-class experiments allow for immediate instructor feedback and discussion while feedback is delayed until the next class meeting with out-side of class experiments run on a computer. Second, in-class experiments run manually allow for considerably more face-to-face interaction between students. These interactions foster a greater sense of community within the class and may lead to higher levels of subsequent (out-side of class) interaction (e.g. study groups).

The concerns regarding online experiments are supported by evidence from the literature. In studies of computer-based or distance learning, students report missing the face-to-face interaction with both their peers and instructor in these alternative learning environments (see for example Staarman et al., 2005; Belcheir and Cucek, 2001; Barkhi and Brozovsky, 2000; Jung et al., 2002). Further, evidence suggests that social interaction can positively influence student learning (Staarman et al., 2005; Jung et al., 2002).

The differences between the two manners of experiment administration are undeniable. It remains a question, however, whether the reduced direct interaction and differences in debriefing between the two experimental media result in any significant differences in student achievement or other outcomes. In this study, we investigate a variety of outcomes from a principles of microeconomics course in order to determine whether there are differences between the two media through which experiments may be administered. We find that courses that employ in-class, hand-run experiments (as opposed to computerized, online experiments) tend to have slightly higher achievement and report more favorable views of the experimental pedagogy. On other dimensions, however, we find no significant difference between the two media including students' perceptions of the interaction level and their overall view of the course, and the degree to which they report working and studying in groups for the course.

II. Data and Empirical Methodology

Study participants were enrolled in one of three sections of microeconomic principles at Baylor University during the spring and fall semesters of 2007.[2] Individual classes were relatively small and ranged in enrollment from 36 to 38 students. All three sections were taught by the same instructor using an identical syllabus – course material coverage and the textbook were uniform in all class sections. The course instructor employed a pedagogy involving the use

[2] Optimally students would be randomly assigned across sections. Such assignment, however, was not possible. At the time of student enrollment students did not know whether they had selected a section with n-class or online experiments. In fact, students did not know whether they had selected into a class employing the experimental pedagogy at all.

of experiments. Over the course of the semester, a total of six experiments were run in each class section. The experiments illustrated a variety of economic concepts including demand and supply in a competitive market, the effects of taxes, price controls and externalities, monopolies and cartels, and the prisoner's dilemma. Table 1 presents a full list of the experiments with a brief description of each. The only difference between the three sections was the medium through which the experiments were administered. Two of the sections, taught during the spring 2007 semester, employed in-class, hand-run experiments while the section taught during the fall of 2007 employed computerized online experiments.[3] With the exception of the Prisoner's Dilemma experiment, in-class experiments were drawn from Bergstrom and Miller's *Experiments with Economic Principles: Microeconomics* (2000). The in-class prisoner's dilemma experiment was from Holt and Capra (2000). The online experiments were developed by Charles Holt and can be found at VeconLab, http://veconlab.econ.virginia.edu/admin.htm.

A. Model of Student Learning

To motivate our empirical work, we use an educational production function approach that is standard in the literature (see, e.g., Siegfried and Fels, 1979). In this approach, the following reduced-form model is specified:

> Student learning = *f*(aptitude; educational background; other student-specific characteristics; educational environment, technology, or teaching methodology; observed and unobserved section-level effects).

[3] All experiments were run during class time. Online experiments were run during the first 30-40 minutes of class and students were allowed to participate in the experiments from any location they wished. For example, students could bring their laptops to class and participate by logging in from there, or they could login from home or a computer lab. After the experiments concluded, students were given 5-10 minutes to arrive in the classroom at which time a debriefing would take place. As a result, class contact time for each group was the same.

Our student learning (achievement) measure takes on two forms: (1) a "gap-closing" measure defined as the difference in post- and pre-course Test of Understanding in College Economics (TUCE) scores expressed as a percentage of the maximum possible point improvement available based on the student's pre-course TUCE score,[4] and (2) course score which is calculated as the percentage of the total possible points earned by the student during the semester. In addition to measures of aptitude (e.g., students' GPA), educational background (e.g., whether a student has taken high school economics) and other student-specific characteristics (e.g., age, gender and ethnicity), we include a dummy variable for online experiments that captures the differential effect, if any, on student learning associated with the online experimental medium (as compared to in-class experiments).

B. Measures of Student Learning and Course Perceptions

Becker et al. (1991) call for multiple measures of student outcomes as indicators of efficacy of teaching approaches. We study the relationship between the medium through which experiments are administered and two general achievement measures. These measures are the Test of Understanding in College Economics (gap-closing measure) and final course score. In addition to these "achievement" measures, we also have a number of measures of student interaction and their perception of the course, generally, and the experiments, more specifically. By employing multiple outcome measures, we increase the likelihood that we will be able to reliably capture any difference between the two experimental media.

The TUCE is a standardized test of economic knowledge aimed at the principles level student. Both microeconomic and macroeconomic versions of the TUCE are available. In this

[4] In other words, the gap-closing measure is defined as (post-course TUCE – pre-course TUCE)/(33 – pre-course TUCE).

study, the microeconomics version of the third edition of the TUCE was administered to all students in the sample on the first and last days of class.[5] This research design allows us to measure differences in learning, or value-added, across students. To provide incentive to exert effort on the pre-course TUCE, students were informed that their performance on the pre-course TUCE would impact their final course grade, but were not told explicitly how their score would be included in this calculation.[6] To induce effort on the post-course TUCE, students were informed that their course grades would be based, in part, on their improvement over their pre-course TUCE score.[7]

In addition to the TUCE, we also measure student achievement with student's final course scores. The final course scores for each student in the study depended upon student performance on eleven problem sets, two midterm exams, and a comprehensive final exam. Final course scores are calculated as a percentage of the total possible points (400) that students earned on all assignments over the semester.

We also consider a number of "non-achievement" outcome measures. These include measures of students' interaction with others in the course and their perceptions regarding a variety of elements of the course. Surveys were administered to students on the first and last days of class to determine their level of "contact" with their classmates. Students were asked to report the number of their classmates with whom they had had any contact, the number that they knew as acquaintances, and the number whom they would identify as friends. Students were also asked to report their perceptions of the interaction level in this course (with both other

[5] The third edition of the TUCE has a total of 33 questions for each version. Instructors are given the option of having students complete either the first 30 questions or all 33 questions. Students in this study were instructed to answer all 33 questions to the best of their ability.

[6] The pre-course TUCE was designed to be a surprise exam; i.e., students were to have had no knowledge of the exam before coming to class because such knowledge could have affected attendance and participation in the study.

[7] During the pre-course TUCE assessment, students were not made aware of this grading method to prevent strategic behavior that could have led to a downward bias of the pre-course TUCE scores. Additionally, the change in TUCE scores entered into grade calculation in an identical fashion across all sections.

students and their professor) as compared to their other courses. Finally, in the exit survey, students were also asked a series of "evaluation" type questions regarding the course to measure their perceptions of the course.

Summary statistics for various outcome and control measures are presented in Tables 2 and 3. The unconditional mean pre-course and post-course TUCE scores do not statistically differ across the two study groups. Similarly the average score earned by each group are not statistically different. Differences do exist across student interaction and perceptions of experiments. These outcomes will be discussed at greater length below.

C. Measuring Inputs in the Educational Production Function

It is essential to control for individual student characteristics. Student achievement depends upon many factors in addition to the pedagogical approach that is the focus of the study. In order to accurately measure the impact of the format in which experiments were administered, our analysis must control for student specific characteristics including gender, age, aptitude, effort, and socio-economic status. We collect a portion of this data through pre- and post-course student surveys including gender, ethnicity, age, employment, high school economics background, and motivations for enrolling in the course. Other measures, however, are more accurately collected directly from student records, including student grade point average and SAT (or ACT) scores.[8] Summary statistics for student characteristics are presented in Table 2. In most respects the students in the two groups are very similar. Average GPA, group gender and ethnicity composition, educational background, employment levels, course load are not

[8] Becker and Powers (2001) argue against using student provided data for aptitude measures due to their unreliability. Further, Maxwell and Lopus (1994) demonstrate that student self-reporting of GPA and SAT scores may suffer from systematic reporting bias. Such nonrandom reporting would produce biased estimates of the relationship between student achievement and educational inputs.

statistically different across the two groups. The only dimension upon which the two groups differ significantly is age where the students exposed to online experiments (19.24 years) were significantly younger than those with in-class experiments (20.98 years).

D. Notes on Estimation Methods

Censoring and selection are each potential issues in our data. For example, each of our dependent variables are potentially subject to censoring problems. The TUCE gap-closing measure can potentially range between $-\infty$ and 1. In our data these values actually range between -0.56 and 0.82. Similarly, final course scores can theoretically range from 0 to 105[9] and in our data range from 14.0 to 102.3. In light of the potential censoring issue, we estimated all of our models using a Tobit estimation procedure. Our results were robust to a variety of specifications.

Selection bias also posed a potential issue in our data. That is, it is possible that the same variables that influence a student's performance with regard to our outcome measures also influence whether or not the student continues in the course such that we are able to construct one or more of our outcome measures for the student. If outcome measures are missing in a systematic manner, failure to control for sample selection will result in biased estimates. In such cases, Becker and Powers (2001) recommend the use of a standard Heckman selection correction to control for any potential nonrandom attrition from the sample. In our sample of 111 students, between 9 and 16 values of the dependent variable are missing (16 TUCE related measure, 9 course scores). To control for any potential selection bias, we employ the Heckman approach and, not surprisingly, find statistically significant selection issues in our data. Importantly,

[9] Extra credit was offered in each of the three sections that allowed for up to 20 additional points. With the extra credit each student could have earned up to 420 points, but their percentage was calculated out of 400 point total resulting in a maximum percentage of 105.

however, estimates of the coefficients on the predictors of student learning, particularly our treatment-group dummy variable, are materially unchanged when controlling for selection.

We find no qualitative (or significant quantitative) differences in our estimates of the effects of participation in the treatment group when we specifically account for censoring and selection issues. Thus, in the results that we present below, we report only our OLS estimates.

III. Results

A. Qualitative Outcomes

Becker (1997) and Becker and Powers (2001) convincingly demonstrate the importance of measuring a wide variety of outcomes when studying the efficacy of different pedagogical approaches. The presence (absence) of evidence regarding student achievement differences between different pedagogical approaches is not sufficient to support the adoption (rejection) of a particular approach. As a result, in addition to our (relatively) objective student achievement measures, we also investigate a variety of other outcomes. We report a set of additional outcomes for our two groups in Table 3.

One argument against computerized experiments is the reduction in face-to-face interaction within the class. In Table 3, we report the number of classmates a student reports having had contact with over the course of the semester.[10] Students in the in-class experiment group report having had contact with an average of 5.34 of their classmates while those in the computerized, online experiment group only report having had contact with 2.79 of their

[10] In both the pre-course and post-course surveys, students are asked to report the number of their classmates with whom they have had direct personal contact. The difference between the two measures is reported in Table 3. While it is possible that this contact came outside of this class, we believe it is more likely that the contact was a direct result of their presence in this course.

classmates – a statistically significant difference.[11] Interestingly, though, when students are asked to compare the level of interaction they had in the course (with either their classmates or the professor) to that of other courses, there is no statistically significant difference between the two groups.

Students are also surveyed regarding the experiments specifically. When asked to rate their agreement with the following statements (on a 5-point scale with 5 = strongly agree, 3 = neither agree nor disagree, and 1 = strongly disagree), "Experiments contributed to my learning of the subject matter in this course" and "Experiments contributed to my overall satisfaction with this course", students in the in-class experiment group reported significantly higher agreement. Further, when asked to list the assignments that they found most valuable, a significantly larger percentage (28%) of the in-class experiment group listed experiments than did the online experiment group (6%).

Other more general evaluations of the course, however, yielded no significant difference between the two groups. For example, students' agreement with statements regarding the instructors' stimulation of their interest in the subject matter or thinking were similar across the groups as was their agreement to the statement "Assignments contributed to my understanding of the course content" and "I learned a great deal from this course".

B. Regression Analysis of Student Learning

We estimate the effect of computerized, online experiments (as compared to traditional, in-class, hand-run experiments) on our measures of student learning for our usable sample of 102

[11] Students were also asked to report the number of acquaintances and friends in their class in both the pre- and post-course surveys. No significant difference in the change in the number of acquaintances or friends was found between the groups. The number of friends in class reported in the pre-course survey did differ significantly between the groups, however, with students in the online experiment group reporting having a greater number of friends in the class.

students.[12] Tables 4 and 5 report estimates for our gap-closing and course score measures respectively. As we discuss in Section II, all estimates are obtained using OLS, but these estimates are robust to corrections for sample selection and censoring.

Table 4 presents estimates on the TUCE gap-closing measure. We present four specifications for the gap-closing measure. In each case, the first specification is a simple difference-in-means estimation. Students in the in-class experiment sections experienced, on average, an improvement of 24.9 percent of the available points on the post-TUCE as measured by the gap-closing measure while students in the treatment group gained 22.6 percent of the available points. The difference is statistically significant. Subsequent specifications add controls for student-specific characteristics including aptitude, gender, age, ethnicity, educational background, and time constraints. Inclusion of these additional controls improves the fit of the model, and the coefficient on the online experiment control remains negative and statistically significant for all of the specifications. Most of the student characteristics are not significant predictors of the gap-closing measure. A student's GPA, however, is consistently a positive and significant predicator of student achievement as measured by the gap-closing measure. The only other statistically significant predictor of the gap-closing measure is whether the student was taking the course for the first time. Students taking the course for the first time showed less improvement than those retaking the course.

Table 5 presents the estimates for the final course score measure. The first specification is again a simple difference-in-means estimation where no significant difference in performance is associated with the experimental medium. Subsequent specifications, (3) and (4) in particular, control for a number of student-specific characteristics. Once student-specific differences are

[12] The usable sample differed for the various outcome measures. The total usable sample was 95 for the TUCE measure and 102 for the final course score.

controlled for, the online experiment group's performance in the course is an average of 3.8 to 4.5 percentage points lower than the in-class experiment group. Not surprisingly student GPA is a significant, positive predictor of student performance while students' age and number of absences are significant, negative predictors. Coefficients on the remaining controls are not statistically significant.

IV. Conclusion

In our study of 111 microeconomics principles students, we investigate whether students who are exposed to the experimental pedagogy using online, computerized experiments experienced any differential in their achievement as compared to students exposed to in-class, hand-run experiments. We find that those students participating in in-class experiments experience higher student achievement than those exposed to online, computerized experiments. We also find that student report greater satisfaction with in-class experiments and find them more valuable to their learning. Students participating in in-class experiments also report having direct contact with a larger number of their classmates than do students participating in online experiments. Yet, when asked to compare the level of classroom interaction in this course with their other courses, there is no significant difference between the two groups. Similarly there is no significant difference between the two groups when students are asked to generally evaluate the course or instructor.

The study provides evidence that students exposed to in-class, hand-run experiments experience higher achievement than students who participate in online experiments. The reason(s) for the difference, however, is unclear. Perhaps students gain more from the experiment debriefings as it is more immediate with in-class experiments where all the participants (including the instructor) can more readily interact. Alternatively, students may experience higher achievement due to indirect effects associated with the social environment in the classroom.[13] The findings may also be a result of the combination of the two influences, or perhaps other factors associated with the different media.

The study also provides evidence that students view in-class experiments more favorably than computerized, online experiments. Humans are social creatures and students may simply derive greater enjoyment from interacting in a face-to-face format than through a computer. Interestingly, though, the difference in the experiment administration does not appear to differentially affect their general views of the course.

The results of this study will help develop a greater understanding of the experimental pedagogy. As experiments are increasingly adopted by instructors (both hand-run, in-class, and computerized, online) it is important to identify the effects of the pedagogy and how the effects (may) differ depending on the medium of experiment administration. Our results suggest that the "devotees" of the traditional hand-run, in-class experiments may be correct when they argue that online experiments are "just not the same" as hand-run experiments and that given the option between the two, hand-run experiments are preferable.

Still, this study is just a first step in expanding our understanding of the experimental pedagogy. Considering the increasing adoption of the pedagogy, future research in this area is necessary. In particular, it would be instructive to replicate the study in other university environments, with other class sizes, with a larger sample, and with a "no experiment" control group.

[13] Students were asked to report the degree to which they studied together for the course. Interestingly, students in the online experiment group actually reported higher rates of studying together. This, however, is likely attributable to the fact that in the pre-course survey students in the online experiment group reported a significantly larger number of friends in their class than did the in-class group.

References

Barkhi, Reza and John Brozovsky. "An Analysis of the Dynamics of a Distance Course" *Journal of Educational Technology Systems*, 28(4), 2000, pp. 311-25.

Becker, William E. "Teaching Economics to Undergraduates." *Journal of Economic Literature*, 35, September 1997, pp. 1347-73.

Becker, W., Highsmith, R., Kennedy, P., and Walstad, W. "An Agenda for Research on Economic Education in Colleges and Universities." *American Economic Review Papers and Proceedings*, 81(2), May 1991, pp. 26-31.

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

Belcheir, Marcia and Mira Cucek. "Student Perceptions of Their Distance Education Courses" *Research Report 2001-04*, Boise State University, Office of Institutional Assessment, 2001.

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

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

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

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

Emerson, Tisha and Beck Taylor. "Interactions Between Personality Type and the Experimental Methods" *Journal of Economic Education*, 38(1), Winter 2007, pp. 18-35.

Holt, Charles and Monica Capra. "Classroom Games: A Prisoner's Dilemma" *Journal of Economic Education*, 31(3), Summer 2000, pp. 229-36.

Jung, Insung, Seonghee Choi, Cheolil Lim and Junghoon Leem. "Effects of Different Types of Interaction on Learning Achievement, Satisfaction and Participation in Web-Based Instruction." *Innovations in Education and Teaching International*, 39(2), 2002, pp. 153-162.

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

Siegfried, John J. and Fels, Rendigs. "Research on Teaching College Economics: A Survey." *Journal of Economic Literature*, 17, September 1979, pp. 923-69.

Staarman, Judith K., Krol, Karen, and Henny van der Meijden. "Peer Interaction in Three Collaborative Learning Environments" *Journal of Classroom Interaction*, 40(1), 2005, pp. 29-39.

TABLE 1. –Experiments

Experiment Name	Brief Description
1. Supply and Demand	Buyers and sellers are given reservation values for a homogeneous product. Trades are carried out in a double oral auction environment. Successive rounds yield prices close to the equilibrium prediction. This experiment serves as the model for many of the future experiments.
2. Minimum Wages	An artificial price floor is imposed in a labor market. Students observe the surpluses generated by the price control.
3. Sales Tax	Tax incidence is examined via a sales tax imposed on the market. Students confirm the allocative equivalence of taxes imposed on either sellers or buyers.
4. Monopolies and Cartels	Firms participate in cartels and are forced to make joint pricing decisions. This restriction is relaxed and reversion to the competitive equilibrium is observed as defection occurs. Emphasis is placed on marginal revenue and marginal cost calculations.
5. Prisoner's Dilemma	Students participate in a prisoner's dilemma where the extent of cooperation is affected by the payoff incentives and by the nature of repeated interaction.
6. Externalities	Students' surpluses are reduced to reflect an external cost of production. Students learn the effectiveness of using Pigouvian taxes or a pollution permit auction to internalize the external cost and correct the market failure.

TABLE 2. – Descriptive Statistics (In-class vs. Online Sections)

Variable	In-Class Experiments Mean (SD)	Online Experiments Mean (SD)
Pre-course TUCE	10.61	10.33
	(0.56)	(0.42)
Post-course TUCE	16.45	15.40
	(0.75)	(0.81)
Course Score	69.40	68.09
	(2.32)	(3.14)
GPA	2.64	2.79
	(0.09)	(0.11)
Male	0.59	0.45
	(0.06)	(0.09)
Nonwhite	0.32	0.23
	(0.06)	(0.08)
Age	20.98	19.24†
	(0.41)	(0.10)
First time taking micro principles	0.84	0.97
	(0.05)	(0.03)
Job	0.35	0.32
	(0.06)	(0.09)
Semester hours completed	29.18	28.21
	(3.41)	(2.70)
Current semester hours	14.95	16.45
	(0.28)	(1.21)
Number of absences	6.53	3.89
	(0.93)	(0.67)
High school course in economics	0.75	0.81
	(0.06)	(0.70)
Number of observations	57	35

† In-class and online groups' means are statistically different at the 5% (two-tailed) significance level or better.

TABLE 3. – Various Outcomes (In-class vs. Online Sections)

Outcomes	In-Class Experiments Mean (SD)	Online Experiments Mean (SD)
Increase in the number of classmates with whom you have had direct contact over the semester	5.34 (0.64)	2.79† (0.63)
Students listed experiments among the most valuable assignments	0.28 (0.06)	0.06† (0.04)
Student perceptions of interaction: (great deal = 10; not at all = 1)		
Level of interaction with classmates in course	5.25 (0.31)	4.59 (0.37)
Level of interaction with professor in course	4.25 (0.31)	4.49 (0.36)
Student evaluations: (strongly agree = 5; agree = 4; neither agree nor disagree = 3; slightly disagree = 2; strongly disagree = 1)		
Instructor stimulated my interest in this subject	3.68 (0.15)	3.94 (0.14)
Instructor stimulated my thinking	4.00 (0.12)	4.03 (0.14)
Assignments contributed to my understanding of the course content	4.18 (0.08)	3.86 (0.19)
Experiments contributed to my learning of the subject matter in this course	3.95 (0.15)	3.09† (0.19)
Experiments contributed to my overall satisfaction with this course	3.95 (0.15)	3.31† (0.20)
I learned a great deal from this course	3.91 (0.11)	3.91 (0.16)
Number of observations	57	35

† In-class and online groups' means are statistically different at the 5% (two-tailed) significance level or better.

TABLE 4. – Gap-Closing Measure (Post-course minus Pre-course TUCE/33 minus Pre-course TUCE)

	Specifications			
Independent variables	(1)	(2)	(3)	(4)
Online experiments	-0.023***	-0.033*	-0.056**	-0.041**
	(0.000)	(0.009)	(0.008)	(0.007)
GPA		0.066***	0.110**	0.075**
		(0.006)	(0.020)	(0.009)
Male			0.008	0.012
			(0.003)	(0.028)
Age			-0.004	-0.006
			(0.002)	(0.004)
Non-white			0.030	0.020
			(0.018)	(0.026)
First time taking micro			-0.023	-0.094*
			(0.050)	(0.029)
Absences			0.007	0.007
			(0.008)	(0.007)
Job				-0.135
				(0.096)
Current semester hours enrolled				-0.008
				(0.004)
High school economics				-0.018
				(0.089)
Constant	0.249***	0.066	-0.007	0.373*
	(0.000)	(0.026)	(0.095)	(0.115)
Observations	91	91	78	78
Adjusted R-squared	0.00	0.04	0.08	0.14

* significant at 10%; ** significant at 5%; *** significant at 1%

NOTES: Robust standard errors are in parentheses and are adjusted for within-section correlation of errors.

TABLE 5. – Course Score (Percent of total available points earned by student)

	Specifications			
Independent variables	(1)	(2)	(3)	(4)
Online experiments	-1.302	-3.296	-4.542**	-3.770*
	(1.830)	(2.430)	(0.520)	(1.289)
GPA		18.083**	15.973**	15.480**
		(2.513)	(2.612)	(1.967)
Male			0.578	-0.231
			(1.344)	(1.745)
Age			-1.161**	-1.213***
			(0.172)	(0.094)
Non-white			0.611	1.075
			(0.400)	(1.037)
First time taking micro			-3.844	-6.453
			(3.484)	(2.216)
Absences			-1.215	-1.375*
			(0.721)	(0.454)
Job				-1.130
				(4.102)
Current semester hours enrolled				-0.548
				(0.264)
High school economics				2.129
				(2.863)
Constant	69.397***	20.983	57.879*	70.538***
	(1.830)	(9.300)	(14.665)	(4.394)
Observations	102	102	82	82
Adjusted R-squared	0.00	0.54	0.62	0.64

* significant at 10%; ** significant at 5%; *** significant at 1%

NOTES: Robust standard errors are in parentheses and are adjusted for within-section correlation of errors.