
A Randomized Assessment of Online Learning

Author(s): William T. Alpert, Kenneth A. Couch and Oskar R. Harmon

Source: *The American Economic Review*, MAY 2016, Vol. 106, No. 5, PAPERS AND PROCEEDINGS OF THE One Hundred Twenty-Eighth Annual Meeting OF THE AMERICAN ECONOMIC ASSOCIATION (MAY 2016), pp. 378-382

Published by: American Economic Association

Stable URL: <https://www.jstor.org/stable/43861048>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



is collaborating with JSTOR to digitize, preserve and extend access to *The American Economic Review*

JSTOR

A Randomized Assessment of Online Learning[†]

By WILLIAM T. ALPERT, KENNETH A. COUCH, AND OSKAR R. HARMON*

This paper contains estimates of the impact of different instructional models that incorporate online course content on learning outcomes of college students of principles of microeconomics using a randomized study design. In the existing literature, there are only three published studies (Figlio, Rush, and Yin 2013; Bowen et al. 2014; and Joyce et al. 2015) that use a random design to explore the impact of online education in a college-length course on learning outcomes. Thus, this study provides an important extension to a literature that is extraordinarily small given the widespread adoption of online teaching and its potential impact on student outcomes at the postsecondary level.

There is a large prior literature addressing the impact of incorporating online content delivery into educational instruction. That literature examines impacts on classes at the primary, secondary, and postsecondary levels. As summarized in a meta-analysis released by the US Department of Education (2010, p. xii), the initial search for articles within this massive literature located 1,132 publications as of 2008. Of those, however, only a handful had a random design and none considered a semester-length college-level course.

Against this backdrop, the Figlio, Rush, and Yin (2013) study of the impact of teaching microeconomics principles in a purely online or face-to-face classroom setting on student learning outcomes was the first randomized study for a semester-length course at the postsecondary level. Students were randomized after enrollment into either a classroom-based or purely

online section. In assessing the difference in outcomes across sections, estimates indicate (Figlio, Rush, and Yin 2013, Table 3) that students attending the live lectures do roughly 3 points better than those in the purely online section on a mid-term exam and about 2.5 points better on their final exam. Thus, the average student performed worse in the online sections.

The study of Joyce et al. (2015) similarly uses a randomized study design to examine different instructional models for principles of microeconomics focusing on a contrast between a face-to-face section that met twice a week versus a blended one that instead met once per week. Both sections had access to online materials. This contrast of two course sections in a blended teaching format that varied by providing more or less face-to-face contact time addresses an important margin of decision making for educational administrators. Across the sections with more or less contact time, the study reports that no significant differences are found for the average final exam score.

The analysis of blended versus purely online delivery of a basic statistics course contained in Bowen et al. (2014), while not directly applicable to instruction of microeconomics, is nonetheless notable within the very small set of studies that employ a randomized research design. The study was conducted across six campuses comparing a machine-based learning model with one hour per week of instructor contact time to classroom courses with about three hours per week of contact time. The study examines a number of learning outcomes including final exam scores and concludes that outcomes across sections are about the same. This comparison of student outcomes associated with a blended course design versus traditional classroom instruction arrives at a similar conclusion to the study by Joyce et al. (2015).

We highlight two important differences in this study compared to the prior literature. First, the random design of this study simultaneously considers three course formats (face-to-face,

*Alpert: University of Connecticut-Stamford, 1 University Place, Stamford, CT 06901 (e-mail: alpert@uconn.edu); Couch: University of Connecticut, 365 Fairfield Way, Storrs, CT 06269 (e-mail: kenneth.couch@uconn.edu); Harmon: University of Connecticut-Stamford, 1 University Place, Stamford, CT 06901 (e-mail: oskar.harmon@uconn.edu).

[†]Go to <http://dx.doi.org/10.1257/aerp20161057> to visit the article page for additional materials and author disclosure statement(s).

blended, and purely online) whereas prior studies provide a single contrast (Figlio, Rush, and Yin 2013, face-to-face versus online; Joyce et al. 2015 and Bowen et al. 2014, face-to-face versus blended). This allows us to consolidate and confirm prior findings within a single study. Second, this study randomly assigned students at the point of the expression of interest in taking the principles course rather than after enrollment (Figlio, Rush, and Yin 2013; Bowen et al. 2014; Joyce et al. 2015). This allows us to examine the potential importance of nonenrollment and failure to complete courses on learning outcomes.

I. The Randomized Trial

The sample was collected from a microeconomics principles course taught over 4 consecutive 16-week semesters at a large public university in the Northeast. Each semester a course of microeconomic principles was listed that had 3 sections capped at 35 students containing one of the instructional models. As an incentive to participate, students were given five points on their term average for the course. Five-hundred nineteen students were randomized across the four semesters.

A. Instructional Formats

The experimental design randomly assigned students to one of three delivery modalities: classroom instruction, blended instruction with some online content and reduced instructor contact, and purely online instruction. The traditional section met weekly for two 75-minute sessions, alternating between a lecture and discussion period. The blended section met weekly with the instructor for a 75-minute discussion period. As a substitute for the lecture period, students were given access to online lecture materials. The online section had class discussion in an online asynchronous forum, and for the lecture period students were given access to the same online lecture materials as students in the blended section. The online materials were developed using best practice standards from the Higher Ed Program Rubric for online education as described on the website of the Quality Matters (2014) organization. For the three arms of the experiment, lectures, discussions, and other instructional content were prepared and delivered by the same instructor. The measure

of learning outcomes examined is a cumulative final exam score.

All three sections were given access to the PowerPoint slides used in the lectures. Here we discuss the additional online lecture materials provided to the online and blended sections. One additional online lecture material was a tape of the live lecture originally given to the face-to-face section. To reduce the incidence of lecture-listening fatigue, the online and blended sections were also given access to a compact, 20-minute version of the lecture that used some of the same PowerPoint slides as the full length one. The compact version is closed-captioned, which is helpful for students for whom English is a second language and hearing-impaired students. Additionally, a version of the shortened lecture closed-captioned in Spanish was provided to the online and blended sections. The closed-captioning is keyword searchable so students can easily locate sections of the lecture for review.

The online and blended sections were also given access to an enhanced PowerPoint version of the lecture with hyperlinks separating it into five manageable chunks and a voice-over of the main points. The slides contain economic diagrams drawn sequentially by user mouse clicks and guided by audio narration. Students are provided with a drawing tool and encouraged to draw the diagram. The script of the audio narration appears in a notes pane and the translate function in PowerPoint will display the script in a language of choice.

B. Online Discussion Forums

As a substitute for the learning from interactions between an instructor and students and between students that can occur in a classroom setting, the best practice recommendation from Quality Matters is to provide an online discussion forum that engages students and fosters a sense of community. In the online section, asynchronous discussion was conducted using two equal-sized Facebook groups. A portion of the course grade depended on the level of participation. Participation included posting (or responding to) questions, posting links to articles or videos helpful for understanding topics in the lecture, and engaging in discussion questions posted by the instructor.

II. Descriptive Statistics and Estimation Results

A. Course Completers

First, we present evidence that across the arms of the experiment (face-to-face, blended, and online) descriptive characteristics of the sample were largely indistinguishable at the point of course completion. Of the randomized students, 323 completed the course.

Online Appendix Tables A.1 and A.2 provide comparisons of average demographic and background academic characteristics for students who completed the course in the face-to-face relative to the purely online (Table A.1) and blended (Table A.2) sections. Across virtually every characteristic, the random design appears to have been successful with mean values for the online and blended sections indistinguishable from the face-to-face sections at conventional levels of significance. The only statistically significant difference (t -statistic = 2.08) indicates that those in the purely online course have taken six additional credit hours in the past. Given the number of comparisons made, this one significant difference is within the range of random error. A control for prior credit hours and the other covariates shown in the descriptive tables are included in our estimates. We conclude that the randomization was successful, most importantly for key characteristics likely to be closely related to subsequent academic outcomes (prior GPA and SAT scores).

We use linear regression for calculating the estimates capturing the difference in outcomes across the face-to-face, online, and blended class modalities for those who completed the course. Panel A of Table 1 contains parameter estimates for differences in scores (out of a possible 100) on a cumulative final exam taken in a similar setting by all students. Column 1 contains estimates not including available covariates. As can be seen there, students in the purely online course score about 4.9 (t -statistic = 3.09) points worse than those in the face-to-face course. In column 2, covariates are added, most related to learning outcomes (prior GPA, prior credits, and SAT scores). With these controls, students in the purely online course are still found to score 5.2 (t -statistic = 3.26) points lower on the final exam. Column 3 includes all available covariates. There, students are still found to score 4.2 (t -statistic = 2.68) points lower in the online sec-

TABLE 1—OLS ESTIMATES FINAL EXAM BY MODALITY

	(1)	(2)	(3)
<i>Panel A. Course completers</i>			
Dummy online = 1	-4.912*** (3.09)	-5.201*** (3.26)	-4.232*** (2.68)
Dummy blended = 1	-1.454 (0.95)	-1.703 (1.12)	-0.996 (0.66)
Observations	323	265	215
<i>Panel B. Adjusting for attrition and noncompletion</i>			
Dummy online = 1	-13.32*** (3.29)	-9.517** (2.20)	-10.28** (2.17)
Dummy blended = 1	-4.261 (1.05)	1.518 (0.35)	-0.121 (0.03)
Observations	519	411	324

Note: Table entries are ordinary least squares (OLS) parameter estimates (t -statistics).

Source: Author calculations

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

tion than the face-to-face variant. The sign of the impact of participating in the blended course is negative but the parameters are not significantly different than zero at conventional levels across the three columns. Online Appendix Table B.1 provides a full set of parameter estimates and diagnostic statistics for the models associated with the entries in panel A of Table 1.

B. Accounting for Differential Attrition

An additional concern with online education is the willingness of students to take a course and the impact of the delivery method on completion. The development of the online tools used in this experiment was consistent with current best practices in course design. Nonetheless, there was clearly greater attrition among those randomly given the opportunity to enroll in the online section.

From the point students received permission to enroll to course completion, potential participation in the face-to-face section declined from 175 to 120 students (30 percent). For the blended course section, the decline was from 172 randomized students to 110 completers (36 percent). The largest decline was observed for the online arm where 172 students were assigned to the course and 93 completed (46 percent).

Descriptive statistics at the point of randomization are provided in online Appendix Tables A.3 and A.4. As seen at course completion, average student characteristics across the course sections are almost identical. Similar descriptive statistics at the point of enrollment are contained in online Appendix Tables A.5 and A.6.

To gauge the potential impact on learning related to differential attrition across the three arms of the experiment, we use the complete sample of those given a permission number along with outcomes for those who completed the course to provide estimates where we assign those who did not complete the course an outcome of zero and recalculate the estimates contained in panel A of Table 1. Those estimates, contained in panel B, indicate that for the online course section, the additional attrition relative to the face-to-face setting amplifies the potential negative impact of the course on the students to be served. As shown in column 1, without including any covariates, the estimated impact for students given the opportunity to enroll in an online course offering has a negative impact of 13.3 (t -statistic = 3.29) points on a scale of 100. Including covariates most related to student learning in column 2, the estimated negative impact is 9.5 (t -statistic = 2.20) points. Including all covariates in column 3, the estimated reduction in the test score is 10.3 (t -statistic = 2.17) points. While all of the estimates for the impact of online course delivery are significantly different than zero at conventional levels of significance, none are for the blended sections. Online Appendix Table B.2 contains a full set of parameter estimates and diagnostic statistics associated with panel B of Table 1.

III. Conclusion

This paper contains results from a randomized study of student outcomes on a cumulative semester exam in a face-to-face course versus two popular variants of online education, a purely online course and a blended model that reduces instructor contact by one-half. Descriptive evidence indicates that randomization was successful at the time permission numbers were assigned, were used, and at course completion. Additionally, across most characteristics, those in the experiment appear similar to students enrolled in a nonexperimental section

of the same course (see online Appendix Table A.7).

Estimates indicate that those who completed the purely online course had learning outcomes that were significantly worse than those in the face-to-face section of the course: this difference is about four to five points or one-half of a letter grade. Differences in outcomes for those who completed the blended relative to the face-to-face course were not significantly different than zero.

Ignoring possible contamination through enrollment outside the experiment, we provide estimates that incorporate the potential negative impact from differential attrition across sections. Those estimates indicate that assignment to the purely online course resulted in a reduction in average scores on a cumulative final exam of about 10 points out of 100—the equivalent of a letter grade. However, no statistically significant findings were found for the blended course format when accounting for attrition.

The learning outcomes on exam scores are not meaningfully different for the average student in the control group (classroom instruction) and blended treatment group for any of the estimates presented, consistent with the results of Bowen et al. (2014) and Joyce et al. (2015). However, exam scores are consistently worse for the online treatment group consistent with findings from Figlio, Rush, and Yin (2013) and earlier studies (e.g., Brown and Liedholm 2002) based on observational data. These results suggest a potentially promising avenue for economizing on teaching resources while maintaining student outcomes is to use blended teaching models of instruction for economics principles that reduce instructor contact time in a classroom setting.

REFERENCES

- Bowen, William G., Matthew M. Chingos, Kelly A. Lack, and Thomas I. Nygren. 2014. "Interactive Learning Online at Public Universities: Evidence from a Six-Campus Randomized Trial." *Journal of Policy Analysis and Management* 33 (1): 94–111.
- Brown, Byron W., and Carl E. Liedholm. 2002. "Can Web Courses Replace the Classroom in Principles of Microeconomics." *American Economic Review* 92 (2): 444–48.
- Figlio, David, Mark Rush, and Lu Yin. 2013. "Is It Live or Is It Internet? Experimental Estimates

of the Effects of Online Instruction on Student Learning.” *Journal of Labor Economics* 31 (4): 763–84.

Joyce, Ted, Sean Crockett, David A. Jaeger, Onur Altindag, and Scott D. O’Connell. 2015. “Does Classroom Time Matter?” *Economics of Education Review* 46: 44–67.

Quality Matters. 2014. “Higher Ed Program

Rubric.” <https://www.qualitymatters.org/rubric> (accessed February 7, 2014).

US Department of Education. 2010. “Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies.” Washington, DC: US Department of Education, Office of Planning, Evaluation, and Policy Development.