

# *A Comparison of Course Delivery Methods: An Exercise in Experimental Economics*

*Roy Howsen and Stephen Lile<sup>1</sup>*

## ABSTRACT

This paper compares student performance in traditional versus online classes using a methodological framework free from problems affecting past studies including self-selection bias. The data employed in this study consists of 1) students' test scores in Principles of Macroeconomics courses, 2) students' academic information collected from the registrar's office, and 3) students' personal information collected from a survey completed by students involved in the study. Employing a maximum likelihood estimation procedure, we find that the average grade for students in the online class is approximately one letter grade lower than the average grade for students in the traditional class, *ceteris paribus*.

## Introduction

Conaway (2002, p.12) notes that Russell (1999) examined over 350 studies that dealt with type of course delivery and concluded that "No matter how it is produced, how it is delivered, whether or not it is interactive, low-tech, or high-tech, students learn equally well with each technology and learn as well as their on campus, face-to-face counterparts". During the past several years, numerous articles have investigated the relationship between course delivery methods and student performance. The two major modes of delivery are the traditional delivery method defined as face-to-face within a classroom setting and the online delivery method defined as a web-based course. Representative prior studies, e.g., Brown and Liedholm (2002); McLaren (2004); Coates et al (2004); Sosin et al (2004); McFarland and Hamilton (2005-06); Stephenson et al (2005); Anstine and Skidmore (2005) have produced confounding empirical results with respect to the effect that traditional and online instruction methods have on student academic performance. Generally, the results suggest that there is either no significant difference between the two modes of delivery or that the traditional delivery has a statistically positive impact on student academic performance relative to online delivery. The purpose of this article is twofold: 1) to note the methodological differences leading to confounding results in prior studies and 2) to provide for a methodological framework free from problems in earlier studies so that one can discern the differential impact of traditional and online delivery instruction on student academic performance.

Prior studies are deficient methodologically in one or more of the following respects: 1) examinations were generated by the instructor of record thereby introducing the possibility that the instructor "taught to the exam", 2) different exams were administered to the traditional and online classes, 3) online students were e-mailed, faxed or mailed their exams, thereby raising the question of sole proprietorship (one author noted that students were on the honor system with respect to not contacting or receiving help from someone else, and that some online students were given the entire weekend to return their examinations), 4) course material, i.e., problems, assignments, notes varied for the online and traditional delivery methods, and 5) the instructors' experience with respect to teaching traditional versus online courses is very seldom addressed.

Without a controlled environment, prior studies could not disentangle the effect of the type of delivery method from other effects embodied within this variable, nor the subsequent effect on student achievement. For instance, if an online class is allowed the entire weekend to finish an examination versus a traditional class that is allowed one hour, one would have no way to control for the effect of the online delivery versus

---

<sup>1</sup> Professors, Department of Economics, Ford College of Business, Western Kentucky University.

the effect of the extended time period. Furthermore, prior studies that either administered different exams or used different instructors for the different delivery methods introduced measurement error into the dependent variable. How this study dealt with these methodological issues and others are explained in the following paragraph.

In designing our study, we wanted to address each of these aforementioned methodological issues in order to provide for a more controlled environment and this was accomplished in the following manner. First, the same instructor taught both the online and the face-to-face classes. Second, the instructor did not generate the examination questions, and did not see the test questions prior to the day the test was administered. Multiple-choice questions were selected by a departmental colleague at random from the test bank that accompanies *Macroeconomics for Today*, 4<sup>th</sup> edition, by Irvin Tucker. Third, the identical set of test questions was given to both the online and to the face-to-face classes. Fourth, online students were required to take each of the four tests on campus on the same day that the traditional class took the tests. Requiring online students to come to campus created a burden on them, but we felt that taking the same test on the same day under the same physical conditions and under the same time constraint was an important methodological feature of this experiment. Fifth, all students used the same textbook (Tucker's *Macroeconomics for Today*), did the same homework problems, and posted comments to the same discussion board topics. The syllabus used in both classes pointed out that performance on the four tests was the major influence on grades, but students in both classes knew that homework and discussion board participation could affect final grades in borderline cases. The instructor posted lecture notes on Blackboard relating to each chapter covered in the Tucker textbook. These notes mirrored the lecture given in the face-to-face class. Although lectures were not recorded for asynchronous access, students in the online class could e-mail or phone the instructor questions about the reading material, lecture notes, homework, or discussion board topics. Finally, unlike some previous studies, the instructor in our study had extensive previous experience teaching the macro-principles course using both face-to-face (34 years) and online (5 years) delivery. Thus, by ferreting out these methodological issues from the delivery method variable, one is left with the unbiased and direct effect of the two various delivery methods on student performance.

### Data and Model

The model in this study is consistent with earlier studies in that a production function approach is employed. Support for this specification can be found in Coates et al (2004) and Brown and Liedholm (2002). Generally, student performance is dependent upon the type of course delivery method, student characteristics, and parental characteristics. The specification of the student performance equation is:

$$SP=(CLASS, GENDER, RACE, AGE, ACT, FATHER, MOTHER, CRHRS, CHE, WORK, MAJOR, MATH, PREVECON) \quad Eq.1$$

where *SP* measures the student's academic performance for the semester; *CLASS* is a dummy variable equal to one if the student was in the online class, zero otherwise; *GENDER* is a dummy variable equal to one if the student is male, zero otherwise; *RACE* is a dummy variable equal to one if the student is white, zero otherwise; *AGE* is the student's age in years; *ACT* is the student's composite act score; *FATHER* is a dummy variable equal to one if the student's Father has a college degree, zero otherwise; *MOTHER* is a dummy variable equal to one if the student's Mother has a college degree, zero otherwise; *CRHRS* is the number of the student's credit hours enrolled during the current semester; *CHE* is the number of cumulative credit hours earned by the student; *WORK* is the number of hours worked per week by the student; *MAJOR* is a dummy variable equal to one if the student has a major within the College of Business, zero otherwise; *MATH* is the student's perception of their math ability; and *PREVECON* is a dummy variable equal to one if the student has previously taken an economic course, zero otherwise.

The focus of this study is the delivery method variable, *CLASS*. If the online delivery method results in increased student performance, the sign on the coefficient of this variable will be positive and statistically significant. Alternatively, if the online delivery method results in decreased student performance, the sign on the coefficient of the *CLASS* variable will be negative and statistically significant.

There are no prior expectations as to the signs on the coefficients for the *GENDER* and *RACE* variables. Prior studies such as Coates (2004) and Brown and Liedholm (2002) have shown these variables have mixed results with respect to their effect on student performance.

Older students are generally more mature and thus more serious about their education. Thus, the sign on the coefficient for the *AGE* variable is expected to be positive. *ACT* scores are generally related to innate ability which in turn leads to higher student performance, Hanushek (1986). Therefore, the coefficient on *ACT* is expected to be positive. Having parents with a college degree would lead one to expect that their children would perform better in school. Thus, the sign on the coefficients *FATHER* and *MOTHER* variables should be positive. Students who are enrolled for three credit hours, i.e. one class, should, ceteris paribus, have more time to devote toward that class thereby leading to increased student performance as opposed to a student who is taking a full course load. Thus, the *CRHRS* variable's coefficient should be negative. *CHE* is a proxy variable for student classification, i.e. freshman, sophomore, junior, senior, and one would expect students with more college experience to perform better. Students who intend to major in the college of business should have an increased desire to outperform non-business majors because admission into the college of business is partly based on their grade point average. Thus, the sign on the coefficient for the *MAJOR* variable should be positive. Students who have taken a prior economics course should perform better versus students who have not had a prior economics class. Thus, the sign on the coefficient for the *PREVECON* variable should be positive. Number of hours worked per week, i.e. *WORK* variable, should have a negative effect on student performance, ceteris paribus.

Since students were not randomly assigned to either the online or to the face-to-face class, there is self selection sample bias. Green (2000) has shown that estimating Equation 1 by ordinary least squares would lead to biased and inconsistent parameters. Such is the case with our study because students could choose to enroll in either the online or in the face-to-face class. Green (2000) equates the self-selection bias to omitted variable bias which tends to overstate the treatment effect, i.e. the effect of the online class. He suggests that the solution to the self selection problem is to develop a second equation in which one models class selection and estimates the class selection equation and the student performance equations by the maximum likelihood procedure. Thus, the class selection equation to be estimated is:

$$\text{CLASS} = (\text{GENDER}, \text{RACE}, \text{AGE}, \text{WORK}, \text{CAMPUS}, \text{FRIEND}, \text{WEB}, \text{INTERNET}, \text{CUMGPA}) \quad \text{Eq.2}$$

where the variables *CLASS*, *GENDER*, *RACE*, *AGE*, and *WORK* are as defined in the above paragraphs. *CAMPUS* is a dummy variable equal to one if the student lives on campus, zero otherwise; *FRIEND* is a dummy variable equal to one if the student of record knows someone who has taken a web course, zero otherwise; *WEB* is a dummy variable equal to one if the student of record has previously taken a web course, zero otherwise; *INTERNET* is the students' perception of their Internet ability on a scale from 1 to 5, i.e. low to high; and *CUMGPA* is defined as the students' cumulative grade point average. It is included in the class selection equation for two reasons: 1) it captures work effort and 2) it assists in identifying the class selection equation. *CUMGPA* is not included in the student performance equation because it is highly correlated with student innate ability, i.e. *ACT*.

A rationale for including *GENDER* is provided by Agarwal and Day (2000) who report that women benefit more than men from courses that utilize technology. Following Coates et al (2004) the *RACE* variable is included. The effect of student age on class selection is unclear. Older students may have higher opportunity cost due to greater human capital and employment experience and thus prefer online classes because such classes interfere less with employment. Alternatively, older students may prefer face-to-face classes due to apprehension about returning to college. Thus, *AGE*, unlike other variables in the class selection equation, is tested using a two-tail test. *WORK* is included on the assumption that an online class is more attractive for students who work because the on-line class lowers their opportunity cost of attending college. A similar argument justifies the inclusion of *CAMPUS*. Students who live off-campus are more likely to take an on-line class because doing so reduces commuting expenses and the hassle of finding a parking spot on campus.

Russell (1999) reports that students tend to report a positive experience with online courses. For this reason the variables *FRIEND* and *WEB* are included because a student is more likely to choose an online class if a friend has taken an online class, or if the student in question has previously done so, assuming that the experience is positive. Students are more likely to select an online class the more comfortable they are using the Internet; therefore, a positive coefficient is expected on the *INTERNET* variable. Finally, students' cumulative grade point average is included because, ceteris paribus, higher GPA in part reflects work effort on the part of students which in turn provides more self confidence and therefore may make

students more confident about being able to handle the independent study associated with taking an on-line class.

As previously mentioned, both Equations 1 and 2 are estimated by maximum likelihood estimation procedure. This procedure results in unbiased and consistent parameter estimates for Equation 1.

The data employed in this study consists of students' test scores and personal information. Data were collected from the registrar's office and from a survey completed by each student. There were a total of 109 students involved in the study: 31 were in the web class and 78 were in the traditional class. Summary data is provided in Table 1.

**Table 1**  
**Means and Standard Deviations for Data by Type of Class**

VARIABLE	TRADITIONAL CLASS	WEB CLASS
Student Performance	74.07 (10.84)	69.84 (10.51)
Gender	.53 (.503)	.35 (.49)
Race	.872 (.34)	.903 (.30)
Age	19.92 (1.70)	22.58 (5.46)
ACT	22.56 (3.72)	22.00 (3.67)
Father	.50 (.503)	.387 (.495)
Mother	.51 (.503)	.42 (.502)
Crhrs	15.44 50.05	14.355 65.44
Che	(21.72)	(29.75)
Work	10.35 (12.64)	23.10 (15.54)
Major	.59 (.495)	.484 (.508)
Math	3.69 (.89)	3.40 (.92)
Prevecon	.705 (.459)	.742 (.445)
Web	.35 (.48)	.45 (.51)
Cumgpa	3.08 (.56)	2.84 (.58)
Campus	.58 (.49)	.19 (.40)
Friend	.89 (.30)	.94 (.25)
Internet	4.28 (.66)	4.45 (.72)
Number	78	31

**Estimation Results**

Table 2 provides estimation results for each independent variable associated with the class selection equation, Equation 2. The first number represents the parameter estimate and the number directly below represents the corresponding T-value. According to the empirical estimates, the coefficient on *GENDER* is negative and statistically significant implying that males are less likely to take an on-line class. The *AGE* variable's coefficient is positive and statistically significant implying that as a student's age increases the student is more likely to take an online class. The *CAMPUS* variable's coefficient is negative and statistically significant suggesting that students who live on campus are less likely to take an online class. The coefficient on the *WORK* variable is statistically significant and positive suggesting that as work hours increase students are more likely to take an online class. Finally, students with higher GPAs are less likely to take an online class. This could suggest that higher achieving students, in order to maintain their grade point average, feel more comfortable in a face-to-face class. The remaining variables, *RACE*, *FRIEND*, *WEB*, and *INTERNET* were not statistically significant at the .10 level.

**Table 2**  
**Maximum Likelihood Estimates for the Class Selection Equation**

VARIABLE	COEFFICIENT (T-STATISTICS)
Intercept	-4.2658 (-1.56)
Gender	-0.7543-a (-2.29)
Race	1.0795 (1.151)
Age	0.1885-a (2.02)
Campus	-0.5890-b (-1.65)
Friend	0.1050 (0.17)
Web	0.0866 (0.26)
Internet	0.1046 (0.45)
Work	0.0218-b (1.79)
Cumgpa	-0.5327-b (-1.76)

*Notes: a and b denote statistical significance at the .05 and .10 levels, respectively.*

Table 3 reports maximum likelihood estimation estimates for each independent variable as specified in student performance equation, Equation 1. The first number represents the parameter estimate and the number directly below represents the corresponding T-value. Five estimations of student performance are provided in Table 1; *AVERAGE* represents students' average numerical grade for the entire semester and *EXAMS 1-4* represent students' numerical score on each semester exam. *Exams 1-4* typically covered four chapters of material.

In order to test for the presence of heteroscedasticity, a variation of the Breusch-Pagan test (1979) was employed. This test examines whether the values of the independent variables are related to the variance of the residuals from the regression equation. To perform the BP test one estimates the residuals from the regression equation, squares the residuals, and regresses the squared residuals on the independent variables. If the F-test confirms that the independent variables are jointly insignificant, the null hypothesis that there

is homoscedasticity cannot be rejected. This test on all five equations resulted in insignificant F-tests, thus homoscedasticity cannot be rejected.

The coefficients on the course delivery variable, *CLASS*, are negative and statistically significant at the .05, .05 and .10 level for exams 1,2 and 3, respectively. The class variable for exam 4 is negative and significant at the .104 level. For these four student performance equations, the coefficients on the *CLASS* variable ranged from -5.74 to -11.07. This finding suggests that students who were enrolled in the online class performed less well than students in the face-to-face class after controlling for self-selection sample bias.

Since a comprehensive final exam was not given, the coefficient on the variable *CLASS* for the *AVERAGE* equation is of particular interest. The coefficient of -8.19 suggests that students' average grade for the semester is almost one letter grade lower for online versus face-to-face students. Alternatively, this suggests that face-to-face delivery, *ceteris paribus*, results in higher student performance relative to online delivery over the entire semester.

The *AGE* and *ACT* variables are positive and statistically significant in the vast majority of the five equations. This suggests that an increase in a student's age and ACT score results in higher student performance, not a surprising result. With respect to the students' grade for the semester, the *RACE* variable was statistically significant and positive implying that white students out performed non-white students. Whereas Coates found no significant difference between males' and females' scores, our empirical results show that females' scores are higher than males' scores in four of the five performance equations. One finding of some interest is that having a previous economics class, *PREVECON*, had a positive and statistically significant effect only on student performance for the first semester exam. This finding is not surprising due to the fact that part of the material on the first semester exam covers material, e.g. supply and demand, that students have had in a previous economics course.

Surprisingly, the impact on student performance of whether or not a parent has a college degree differs depending on gender of the parent. *MOTHER* is not statistically significant in any of the five equations whereas *FATHER* is statistically significant in three of the five equations. The variables *CRHRS*, *CHE*, and *MAJOR* are, for the most part, statistically insignificant in explaining student performance. The variable *WORK* surprisingly has a positive and significant effect in student performance exams 1 and 2, but was statistically insignificant in the remaining three student performance exams. Students' math ability is positive and statistically significant for explaining average grade for the semester. It is not surprising that one's math ability has a positive impact on average performance for the semester. A correlation matrix of the independent variables suggests that multicollinearity is not a serious problem. Age and current credit hours taken exhibited the highest correlation which was a negative .5.

**Table 3**  
**Maximum Likelihood Estimates for the Student Performance Equations**

VARIABLE	AVERAGE	EXAM 1	EXAM 2	EXAM 3	EXAM 4
Intercept	-21.729 (-1.31)	-31.042 (-1.25)	-10.393 (-0.48)	-18.026 (-0.82)	-27.454 (-1.62)
Class	-8.192-a (-3.29)	-9.675-a (-2.60)	-11.078-a (-3.83)	-6.268-b (-1.89)	-5.748 (-1.62)
Gender	-6.223-a (-2.75)	-9.170-a (-2.71)	-6.300-a (-2.11)	-3.387 (-1.12)	-6.040-b (-1.87)
Race	9.872-a (2.56)	9.112 (1.58)	12.655-a (2.49)	7.345 (1.43)	10.377-b (1.88)
Age	2.138-a (3.80)	2.033-a (2.41)	2.520-a (3.40)	1.992-a (2.66)	2.005-a (2.49)
ACT	0.982-a (3.45)	1.418-a (3.33)	0.600 (1.60)	1.026-a (2.71)	0.889-a (2.19)
Father	3.395-b	2.180	1.140	4.709-a	5.281-a

	(1.92)	(0.83)	(0.49)	(2.00)	(2.09)
Mother	1.607	3.924	-0.875	3.016	0.362
	(0.94)	(1.53)	(-0.39)	(1.32)	(0.15)
Crhrs	-0.015	0.117	-0.668	-0.589	1.079-a
	(-0.04)	(0.22)	(-1.41)	(-1.23)	(2.10)
Che	0.066-b	0.090	0.094-b	0.025	0.056
	(1.83)	(1.67)	(1.96)	(0.52)	(1.08)
Work	0.143	0.326-a	0.313-a	-0.186	0.120
	(1.59)	(2.26)	(2.46)	(-1.45)	(0.87)
Major	-0.188	-6.590-a	0.683	1.772	3.379
	(-0.10)	(-2.35)	(0.28)	(0.71)	(1.26)
Math	2.495-a	1.690	1.790	5.038-a	1.462
	(2.63)	(1.19)	(1.44)	(3.99)	(1.08)
Prevecon	-0.440	4.910-b	-3.302	0.777	-74.135
	(-0.23)	(1.74)	(-1.32)	(0.31)	(-1.53)
R-Square	.45	.39	.19	.46	.27

Note: a and b denote statistical significance at the .05 and .10 levels, respectively.

### Summary and Conclusions

Experimental economics is fraught with many difficulties, notably the difficulty of holding constant other relevant variables. Moreover, a self-selection bias is introduced in cases where experiments cannot be conducted by assigning participants randomly to different treatments. Our study represents an improvement over previous studies in methodology and, unlike most studies, corrects for self-selection bias. Our evidence from comparing test score performance in the Principles of Macroeconomics courses at Western Kentucky University suggests that online delivery results in substantially lower average test scores, ceteris paribus. This finding does not, however, imply that online delivery of the Principles of Macroeconomics course should be abandoned. Such a conclusion must await further study of the costs and benefits for both online students and institutions providing online classes.

In summary our findings add to the body of knowledge with respect to Thomas Russell's (1999) text, *No Significant Difference Phenomenon* in that our study finds a significant difference in the case of Principles of Macroeconomics when other relevant variables are controlled for and correction is made for self-selection bias.

### References

- Agarwal, Rajshree and A. Edward Day, (1998), "The Impact of the Internet on Economic Education," *Journal of Economic Education*, 29, 2.
- Anstine, Jeff and Mark Skidmore, (2005), "A Small Sample Study of Traditional and Online Courses with Sample Selection Adjustment," *Journal of Economic Education Research*.
- Barnow, C., G. Cain, and A. Goldberger, (1981), "Issues in the Analysis of Selection Bias," *Evaluation Studies Annual Review*, Beverly Hills: Sage Publications.
- Breusch, T. S., and A. R. Pagan, (1979), "A Simple Test For Heteroscedasticity And Random Coefficient Variation," *Econometrica*, 47, 5, 1287-1294.

Brown, Byron and Carl Liedholm (2002), "Can Web Courses Replace the Classroom in Principles of Microeconomics?" *American Economic Review*, 92, 2, May.

Coates, D., B. Humphreys, J. Kane, and M. Vachris, (2004), "No Significant Difference Between Face-To-Face And Online Instruction: Evidence From Principles Of Economics," *Economics of Education Review*, 23.

Conaway, Carrie, (2002), "Virtual University: Is Online Learning Changing Higher Education?" *Regional Review* (Federal Reserve Bank of Boston).

Greene, W. H. *Econometric Analysis*, (2000), 4<sup>th</sup> ed., Prentice Hall, New Jersey.

McFarland, Daniel and Diane Hamilton, (2005-06), "Factors Affecting Student Performance and Satisfaction: Online Versus Traditional Course Delivery," *Journal of Computer Information Systems*, Winter.

McLaren, Constance, (2004), "A Comparison of Student Persistence and Performance in Online and Classroom Business Statistics Experiences," *Decision Sciences of Innovative Education*, 2, 1.

Russell, Thomas, (1999), *The No Significant Difference Phenomenon*, Raleigh, N.C.: North Carolina State University.

Sosin, Kim, Betty Blecha, Rajshree Agarwal, Robin Bartlett and Joseph Daniel, (2004), "Efficiency in the Use of Technology in Economic Education," *AEA Papers and Proceedings*, May.

Stephenson, Kurt, Anya McGuirk, Tricia Zeh and Dixie Reaves, (2005), "Comparisons of the Educational Value of Distance Delivered versus Traditional .

Classroom Instruction in Introductory Agricultural Economics," *Review of Agricultural Economics*, 27, no. 4.