Term paper quality of online vs. traditional students

Stephen Hayward
West Texas A&M University

Rex Pjesky
West Texas A&M University

ABSTRACT

This paper uses a blind grading process to test if the performance of online students were different from traditional students using a term paper from an economics graduate course. Consistent with the literature, no significant difference was found between the scores of online students and those of traditional students. Also consistent with the literature, the only significant factors that influence student scores are variables that proxy for effort.

Keywords: assessment, economics, graduate, online, traditional

INTRODUCTION

While the technology used and the format certainly has changed, distance learning has been a viable choice for students in the entire industrial era. The earliest form of what is now called distance learning were probably correspondence courses transmitted between teacher and pupil by the postal service. This slow means of communication between teacher and pupil has transformed into an environment where teacher and student can communicate costlessly with each other over virtually any distance instantly. While in the past, students doing distance learning would be isolated from each other, now they are not. Students in distance learning classes frequently communicate with one another as easily as they communicate with the professor (Gerlich, Mills and Sollosy, 2009).

This rapid change, which has mostly occurred within the last generation, has raised many new and old questions about how people learn and what the best relationship between professor and student is. To any engaged professor, and to any concerned student, an important question should be if the quality of the learning is equal across formats. It is possible that there is an ineffable quality to the teacher/student relationship and that technology can never replace this relationship. In contrast, maybe technology can make possible superior associations between teacher and student, and the academy is irresponsible if it does not embrace all the possibilities that technology presents (Bower, 2001).

To complicate the question, much of the demand for distance education does not stem from concerns of quality, but rather of convenience. Many students opt for online courses simply because their work or family obligations make taking classes in traditional formats difficult. Conversely, many students elect to pursue higher education in the traditional format because of the social experiences that being on campus are valuable to them. Regardless of the answers to these questions, it is important to make sure that the academy provides the best possible learning experience to students no matter what the mode of delivery is, online or traditional (Bower, 2001).

The purpose of this study is to estimate the difference in learning outcomes, as measured on the performance on a written paper, between students who enrolled in an online section of a class and students who enrolled in a traditional face-to-face section of the same class. The study's design assumes that variations in performance might come from variation in academic setting, in demographic characteristics, in academic ability, and in effort. Many, including Dutton, Dutton and Perry (2002) and Yukselturk and Bulut (2007) have taken this basic approach to assess the differences in online and in class modes of instruction.

The rest of the paper is organized as follows: First, there is an overview of the relevant literature to provide a context as to why this study is important and to provide a justification of the methods of this study. Then there is a discussion of some of the theoretical issues regarding the estimation of the data collected. Third, there is a description of the data and how it was collected. Fourth, an empirical model is developed to make inferences about the significance of the impact that mode of instruction and the other variables in the study have on the performance of the written paper. The final section will conclude the report and discuss implications.

LITERATURE REVIEW

There is a vast literature on distance learning. Much of this literature focuses on quality. The issue always has been that of quality. That is, those who advocate distance learning as a

viable alternative to the traditional classroom believe that the experience should be at least as academically rich as the experience that a student in a traditional classroom receives (Bower, 2001).

There are those that believe that the advantages of distance courses are so vast that the academy is doing a disservice to students in traditional course by not incorporating many of the tools and strategies used in the best practices of distance learning professors. Indeed, Turoff (1999) reports that if a professor takes an active role in facilitating the learning of the student (being the "guide on the side") and designs the course around a collaborative learning environment distance learning classes will have superior outcomes.

Yet most professors are trained in "hand to hand" teaching and indeed many of them were attracted to academics in the first place because they enjoyed direct interaction with students. The reality might be that the faculty member as the "sage on the stage." (Sanderson, Phua and Herda, 2000) The face-to-face instructor will use his/her intuition to gauge student reactions to create an effective classroom environment for learning. The technology of distance learning interrupts that feedback and perhaps denies a critical part of the learning process. Therefore, faculty might not just be luddites refusing to change for their own selfish reasons. Rather, they are simply refusing to follow the latest trend before they are presented with clearer evidence that the new technology helps students learn (Bower, 2001). Furthermore, it is not clear that, even though delivery options and technology are becoming more sophisticated and less costly, professors who deliver the courses or the information technology experts who deliver the machinery understand how people actually learn well enough to design a very effective learning environment that separates student from teacher (Greenagle, 2002).

Clearly, the question of quality in distance and online delivery of course material is an empirical one that can be tested. Terry (2007) used a large cross section of graduate students in business at a regional university to test for differences in learning (as measured by performance on class assignments) in campus classes, hybrid (those that use both online and in class instruction methods) classes, and online classes. Terry found that there was no significant difference in performance by students in campus and hybrid classes but students in online classes failed to perform as well on class assignments and did not meet course and program objectives as well as those students who had direct contact with professors.

DATA COLLECTION

The data for this report was collected from two sections of ECON 6306, a master's level macroeconomics course. One of the sections was purely online while the other was purely a "face to face" class. The classes were designed to minimize the difference between the sections.

To try to achieve equality of design between the online and traditional sections, several steps were taken. First, the classes had common objectives, textbooks and exams. The exams in the traditional class, like those in the online class, were "take home" tests consisting of essay and problem solving questions.

Both classes had weekly assignments that consisted of questions written by the professor that paced the students through the book. Students in the traditional section were held accountable for their performance on these questions by oral quizzing by the professor during class. Students in the online had the same questions, but they had to turn in written answers to the questions. Any relevant readings were distributed to both classes, and any substantial

discussions that occurred in the traditional class were also posted in message boards for the online class.

To maintain the integrity of the experiment, both online and traditional sections were given the same instructions, given the same background information on the topic. Students were to explore if the recent rounds of quantitative easing monetary policy by the world's central banks represented new, innovative policy actions by the central banks or were simply rehashes of old, traditional monetary policy. Economists were debating this topic in the press and in the blogosphere, so students in both sections were pointed to the same sources to use as references.

Finally, the papers were graded by the non-teaching coauthor of this study. The teaching coauthor numbered the papers so the grader would not know the name of the student or if the student was enrolled in the online section or the traditional section. This blind grading process was necessary since grading of paper is subjective and hence might be subjective to bias if the grader knew the name of the student or their online status.

Each paper on Quantitative Easing was given two grades, a traditional 100-point score and a separate critical writing score of ten points. The traditional score was broken down by content, summary, length and references with each category being awarded 40, 20, 20 and 20 points respectively. Since Quantitative Easing was a recent, hotly debated subject among economists as to its ability to work as a beneficial macroeconomic tool it allowed each student to take a position on whether it would work or not work thus giving them an opportunity to use critical writing. The ten points awarded for critical writing were divided into five points for stating a position clearly and five points if the position was defended.

METHODOGY

Davisson and Bonello (1976) proposed that student learning can be treated like a production function in economics where production (i.e. student learning or performance) is a function of class format or mode of instruction, the ability of the student, the effort of the student, and the demographic characteristics of the student. That equation is:

$$Y_i = F(D_i, A_i, E_i, V_i) \quad (1)$$

Where

Y = performance on an assignment or learning outcome;
D = an indicator variable identifying different modes of instruction;
A = a student's inherent academic ability;
E = a student's effort;
V = a vector of demographic characteristics
i = an individual student in a cross section of students

Becker (1983) suggests a simple yet effective empirical model to estimate the relationship between student learning and vector of factors that might influence student learning. Equation (1) can be specified into an equation that can be estimated using data that is of particular interest to a specific research question. In the case of this report, we will estimate the model in equation (1) using the following equation:

$$Paper_i = \beta_0 + Online_i\beta_1 + GMAT_i\beta_2 + GPA_i\beta_3 + Overall_i\beta_4 + Citizen_i\beta_5 + MBA_i\beta_6 + Gender_i\beta_7 + \varepsilon_i \quad (2)$$

Where
Paper = grade on the Quantitative Easing assignment;
Online = 1 is student was in the online section, 0 otherwise;
GPA = the student's cumulative graduate grade point average;
GMAT = the student's score on the Graduate Management Aptitude Test;
Overall = the student's final grade in the course exclusive of the paper grade;
Citizen = 1 if the student is a US citizen, 0 otherwise;
MBA = 1 if the student is an Master's in Business Administration student, 0 if the student is a Master's in Economics and Finance student;
Gender = 1 is a student is female, 0 is a student is male;
i = 44 students.

Equation (2) was estimated using a Tobit maximum likelihood technique. The Tobit estimator was selected instead of ordinary least squares since there was an upper bound on the dependent variable. The Tobit model presumes that if there is a discrete limit on the values of the right hand side variable, then some of the predicted values of the right hand side generated by the regression process may not exist. For example, it may be that one would expect students with a 3.5 GPA to score a 100 percent on an assessment test. This would imply that students with a higher GPA should score more than 100 percent of the assessment test. This, of course is not possible given that 100 percent correct is a practical upper limit on the assessment. Under these conditions, using OLS would not produce consistent estimates of the parameters. (Amemiya, 1985; Long, 1997) Five students in the sample of 44 received a grade of 100 on their paper, so a Tobit procedure is appropriate.

Table (1) presents the summary statistics. Fifty-two percent of the students were in the online class, 36% were US citizens, 64% were MBA students and 43 percent were female. The average score on the paper was 84%. The average grade of the student's tests and participation (That is, the average student grade exclusive of the paper) was 91%. The average GMAT was 484 and the average GPA was 3.64.

Discussion

The results of the estimation of Equation (2) are presented in Table (2). The most important result is the coefficient on the "ONLINE" variable. The coefficient is 5.82, which suggests that on average online students scored about 5.82% higher than traditional students do on the paper, holding the other factors constant. This coefficient however is not significant. The t-stat is 0.98, which suggests there is a high likelihood that this coefficient could happen by random chance in this sample.

Similarly, the coefficients on "GMAT," "CITIZEN," "MBA" and "GENDER" were insignificant. This is consistent with the literature that students' demographic characteristics generally do not affect student performance. The women in the sample scored 8.22 percentage points higher than the males in the sample, but with a t-stat of 1.87, the difference was not significant at the 5% level. There is little evidence in the literature that a student's innate ability as measured by achievement tests (the GMAT) have any predictive power on a student's performance on an individual class or assignment. Also, there is no evidence in the literature that a student's citizenship has predictive power on a student's performance on an individual class assignment. (Gerlich, et. al., 2009) A priori, the author's did not expect there to be any

difference in the performance of masters of business administration students and masters of finance and economics students on this assignment. The estimation of Equation (2) confirmed this expectation for this sample.

The coefficient on "GPA" was 16.44 indicating that students with higher grade point averages did score higher on the paper assignment. The t-stat was significant with a p-value of 0.038, which indicated a high level of confidence that a student's GPA was a significant predictor of performance on the paper. Again, as reported in Gerlich et. al., this is consistent with the literature.

The most interesting and surprising result was the coefficient on "OVERALL." The variable "OVERALL" was the student's grade in the class exclusive of the paper grade (that is, the student's grade on participation and tests.) The coefficient on "OVERALL" was negative and highly significant with a t-stat of -3.58, which implies that a higher grade in the course was associated with a lower grade on the paper. A priori, the authors expected the coefficient on this variable to be positive suggesting that a high grade on the paper would be associated with a high grade in the class.

There are several explanations for this result. Since different individuals graded the papers and the rest of the classes' assignments, it is possible that this result was driven by difference in grading methodology. This is highly unlikely. A second possibility is that spurious correlations among the variables could have driven the results. Finally, and this is the most likely explanation, students who had done well on the midterm exams and in class participation did not put in the effort on the paper assignment as students who had done poorly on the midterm and other assignments. This explanation is consistent with the notion that effort is a strong predictor of student success.

The lesson that can be drawn from this is that if this study and much of the literature is correct, student learning comes from student effort. This is not surprising. Professors need to focus much of their efforts in designing assignments in a way that aligns student effort with the learning objectives and goals of courses and programs to maximize student learning.

CONCLUSION

The purpose of this study was to test if innate academic ability, as measured by GMAT scores, academic effort, as measured by GPA, or academic environment, as indicated by choice to take a "face to face" class or online class had an impact on a student's performance on a short research paper. A large body of literature has found that a student's effort is the most significant predictor of a student' success in a particular class or assignment (Gerlich et. al., 2009).

The results of this study indicate that GPA and the grade a student received in the other parts of the course is the only statistically significant predictor of a student's grade on an individual assignment. This reinforces the conclusions of Gerlich, et. al. which reported that a large set of demographic and course design properties did not predict student performance but that GPA, which proxies for effort did.

A lesson that one can draw from the literature in general, and from this study in particular, is that professors need to design assignments and courses in a way that gives students the incentive to increase their effort regardless of class format. Professors need to carefully design assignments to make sure that student effort is spent focusing on the learning outcomes of individual courses and programs.

REFERENCES

Amemiya, Takeshi. (1985). *Advanced Econometrics*. Oxford: Basil Blackwell.

Becker, W. (1983). Economic education research: New direction on theoretical model building. *Journal of Economic Education, 14*(1), 4-9.

Bower, B. (2001). Distance education: Facing the faculty challenge. *Online Journal of Distance Learning Administration, 4*(2), http://www.westga.edu/~distance/ojdla/summer42/summer42.html.

Davisson, W., & Bonello, F. (1976). *Computer assisted instruction in economics education*. South Bend, IN: University of Notre Dame Press.

Dutton, J., Dutton, M., & Perry, J. (2002). How do online students differ from lecture students? *Journal for Asynchronous Learning Networks (JALN), 6*(1), 1-20.

Gerlich, R. N., Mills, L. H., & Sollosy, M. (2009). An Evlauation of Predictors of Achievement of Selected Outcomes in a Self Paced Online Course. *Research in Higher Education Journal*. (4)

Greenagle, F. L. (2002). The illusion of e-learning: Why are we missing out on the promise of technology? http://www.league.org/publicantion/whitepapers/0802.html

Long, J. S. (1997). *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage Publicantions.

Sanderson, A., Phua, V.C., & Herda, D. (2000). *The American faculty poll.* Chicago: National Opinion Research Center.

Terry, N. (2007). Assessing the difference in learning outcomes for campus, online and ybrid instruction modes for MBA courses. *The Journal of Education for Business, 82*(4), 220-225.

Turoff, M. (1999, November). An end to student segregation: No more separation between distance learning and regular courses. Paper presented at the Summary of Invited Plenary for Telelearning, 99, Montreal Canada.

Yukselturk, E., & Bulut, S. (2007). Predictors for student success in an online course. *Educational Technology & Society, 10*(2), 71-83.

APPENDIX

Table 1: Summary Statistics

Variable	Observations	Mean	Std. Dev.	Min.	Max
Paper	44	83.50	14.67	40	100
Online	44	0.52	0.51	0	1
Citizen	44	0.36	0.49	0	1
MBA	44	0.64	0.49	0	1
GMAT	44	484.31	91.61	340	680
GPA	44	3.64	0.36	2.86	4.00
Overall	44	91.33	9.52	50.0	100
Gender	44	0.43	0.50	0	1

Table (2): Results of the Tobit Estimation	
Variable	**Coefficient (t-statistic) [p-value]**
ONLINE	5.82 (0.98) [0.33]
GMAT	-0.0063 -0.26 [0.80]
GPA	16.44 2.15 [0.038]
OVERALL	-0.81 -3.58 [0.00]
CITIZEN	6.92 0.96 [0.34]
MBA	0.64 0.14 [0.89]
GENDER	8.22 (1.87) [0.07]
Constant	91.20 3.42 [0.00]
Observations	44
Psudeo R2	0.05