March 2006

Do Proctored Exams Matter In Online Classes?

Oskar R. Harmon
University of Connecticut

James Lambrinos
Union University

University of Connecticut

Department of Economics Working Paper Series

Do Proctored Exams Matter In Online Classes?

Oskar R. Harmon
University of Connecticut

James Lambrinos
Union University

Working Paper 2006-08

March 2006

---

341 Mansfield Road, Unit 1063
Storrs, CT 06269–1063
Phone: (860) 486–3022
Fax: (860) 486–4463
http://www.econ.uconn.edu/

Abstract

Does the format of assessment (proctored or un-proctored exams) affect test scores in online principles of economics classes? This study uses data from two courses of principles of economics taught by the same instructor to gain some insight into this issue. When final exam scores are regressed against human capital factors, the R-squared statistic is 61.6

Journal of Economic Literature Classification: A2

Keywords: economic education, distance education, assessments, online instruction, pedagogy

INTRODUCTION

Online offerings of economics classes have experienced a recent growth surge. (Sosin 1997) in fall 1997 surveyed 986 economics departments at post-secondary institutions and received 325 completed surveys for a response rate of 33%. Of the respondents only 24 institutions offered a total of 40 online courses. (Coates 2001) conducted a similar survey just three years later of approximately 750 higher education institutions and received approximately 260 completed surveys for a response rate of 35%. Of the respondents 120 institutions offered 189 economics courses online. A comparison of the two surveys shows that in the three year interval the number of institutions offering online economics courses increased by 400% and the number of these courses increased by 373%.

With the rise in popularity of online courses there has been increased discussion about the issue of whether or not assessments in online courses should be proctored. The issue arises because in the absence of the ID confirmation afforded by a proctored exam, it is impossible to know whether the registered student or a substitute has taken the assessment, or if students worked collaboratively on the exam. A review of the literature on this issue reveals that educators have differing views about the best practice (Serwatka 2003), (Deal 2002), (Shuey 2002), (Taylor 2002), (Rovai 2001), (Young 2001), (Edlin 2000), (Liefert 2000), (Kushner 1999), and (Vachris 1999). One view reasons that commercial testing centers and alternative non-profit testing collaborations offer a feasible option for the proctored format. An alternative view reasons that with appropriate adjustments in the un-proctored online format (e.g. randomized questions from a large pool of test questions, open book testing, etc.) the probability of cheating in the two formats can be brought to a similar level. The literature on the extent of online cheating consists of anecdotal evidence, for example (Carnevale 1999), and (Vachris 1999). Examples of courses that employ un-proctored online assessments in online economics courses are (Coates 2004), and (Vachris 1999) and (Navarro 2000) is an example using proctored assessments,

The purpose of this study is to begin to fill the gap of empirical research in the literature on online cheating. This study uses data from two courses, an online class in Principles of Macroeconomics taught in Summer 2004 and another taught in Summer 2005 for the Online Division of the School of Continuing Studies at the University of Connecticut.

The courses, though offered a year apart were almost identical in structure and content. Each was offered in the course management software WebCT. Each had three hour long exams weighted 18% of the course grade (a total of 54%), required participation in a discussion bulletin board for each chapter weighted 18%, and a cumulative 90 minute final exam weighted 28%. The required readings consisted of chapters in a standard Principles of Macroeconomics textbook. The online instructional materials included, PowerPoint presentations augmented with audio sound files, online practice problems in Excel spreadsheets, and readings from the online edition of the Wall Street Journal as background for participation in twice weekly instructor moderated online discussions. The sole significant difference between the two courses was that in Summer 2004 the final exam was un-proctored and in Summer 2005 the final exam was proctored.

DATA

The variables used in the study and their definitions are shown in Table 1, and the descriptive statistics are shown in Table 2. The source of our data is exam scores in the course and University records.

Table 1	
Definitions of Variables	
Variable	Definition
Age	Age in years
Econ_Major=1	1 if an Economics or Business Major, 0 otherwise
exam1	Score on Exam 1
exam2	Score on Exam 2
exam3	Score on Exam 3
exam4	Score on Final Exam
GPA	Cumulative GPA at end of semester
Grade Level	1 if Freshman, 2 if Sophmore, 3 if Junior, 4 if Senior

A test of the difference between the means of the variables for each course is reported in column 4 of Table 2. The attributes of the students (measured by GPA, Grade_level, Econ Major, and Age) are similar between the courses. Though the mean GPA and Grade level are marginally higher for the Summer 2005 course, and the percent economics/business majors and the average age is marginally higher for the Summer 2004 course, these differences are not statistically significant at the 10% level. On balance, the two sections have approximately the same average level of human capital endowments, though perhaps there is a marginal bias in favor of the Summer 2005 section if we give an edge to the relatively higher GPA and Grade level.

A comparison of the means on the final exam, however, shows that the mean score for the proctored final exam in the Summer 2005 course is about 5 points higher than for the un-proctored final exam in the Summer 2004 course, and the difference is statistically significantly at the 10% level. This difference occurs notwithstanding the rough equivalence of the human capital endowments between the two sections. A comparison of the first three exams, which are un-proctored in both sections, reflects that in the first two exams the mean exam scores are not statistically different at the 10% level of significance, but for the third exam the mean is 10 points higher for the Summer 2005 course and this difference is statistically significant at the 5% level.

Table 2				
Descriptive Statistics				
	Summer 2004	Summer 2005	t_test of difference between 2004 and 2005 Means	
Variable	Mean	Mean		
(1)	(2)	(3)	(4)	
	Mean	Mean		
Exam 1	65.50	70.40	-1.12	
	(14.50)	(15.62)		
Exam 2	85.50	84.57	0.26	
	(12.35)	(13.15)		
Exam 3	67.38	78.09	-3.13	**
	(12.00)	12.55		
Final Exam	72.30	77.16	-1.54	*
	(13.17)	(10.33)		
GPA	2.93	2.99	-0.41	
	(0.46)	(0.56)		
Grade Level	2.24	2.47	-1.10	
	(0.70)	(0.79)		
Econ_Major=1	0.29	0.21	0.68	
	(0.46)	(0.41)		
Age	20.71	20.51	0.20	
	(4.47)	(3.03)		
Sample Size	21	35		
* denotes significant at the .10 level				
** denotes significant at the .05 level				
Exam has 1 missing observation for Summer 2004, and Exam2 has 1 missing observation for Summer 2005				

THE MODEL

The model used was determined by data availability and past research studies (Coates 2004) and (Brown 2002). It is:

$$Exam = b_0 + b_1\ GPA + b_2\ D\_Grade\_Level + b_3\ Econ\_Major + b_4\ Age + u_i$$

The variable GPA and Grade_Level are expected to have a positive sign. Econ_Major is expected to have a positive sign as majors in the discipline of the course are expected to have greater motivation to perform well. Age is also expected to have a positive sign as older students tend to exercise greater responsibility toward academic achievement. The term u represents a random error term.

RESULTS

The results of the 8 OLS regressions (one for each of 4 exams in the two courses) are reported in Table 3. There are two notable observations to be made regarding the details of the regressions in Table 3. GPA is the only variable consistently statistically significant across most of the the 4 exams and the 2 courses. Econ_Major is statistically significant at the 10% level in the two models with the highest R squared (columns 1 and 8).

Table 3

Determinants of Exam Score

	Summer 2004				Summer 2005			
	Exam 1	Exam 2	Exam 3	Final Exam	Exam 1	Exam 2	Exam 3	Final Exam
Variable	Parameter Estimate	Parameter Estimate	Parameter Estimate	Parameter Estimate	Parameter Estimate	Parameter Estimate	Parameter Estimate	Parameter Estimate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	37.60	63.08 ***	62.49 **	63.77 **	9.66	45.21 **	32.01 **	32.78 ***
	(25.04)	(23.28)	(25.62)	(28.97)	(22.81)	(22.37)	(17.91)	(11.58)
GPA	11.72 **	10.94 **	6.77	7.71	16.44 ***	10.45 **	12.95 ***	13.11 ***
	(6.92)	(6.44)	(7.08)	(8.01)	(4.44)	(4.46)	(3.49)	(2.25)
Grade Level	17.34 **	4.33	0.41	-1.39	2.55	-3.33	3.31	1.01
	(8.11)	(5.78)	(6.36)	(7.20)	(3.54)	(3.35)	(2.78)	(1.80)
Econ_Major=1	-12.54 *	-6.94	-4.45	-0.82	-9.86	-1.96	1.27	5.00 *
	(8.36)	(6.05)	(6.65)	(7.52)	(6.25) **	(5.88)	(4.91)	(3.17)
Age	-2.10 **	-0.86	-0.70	-0.52	0.35	0.80	-0.06	0.79
	(1.03)	(0.83)	(0.91)	(1.03)	(0.94)	(0.89)	(0.74)	(0.48)
R-Squared	0.4460	0.3547	0.0173	0.1219	0.3483	0.1894	0.3781	0.6160
F-Ratio	3.27 **	2.20	0.84	0.56	3.87 ***	1.64 ***	4.41 ***	11.63 ***
N	20	21	21	21	34	33	34	34

Standard errors are in parentheses below the parameter estimate.

* denotes significant at the .10 level for one-tail test

** denotes significant at the .05 level for one-tail test

*** denotes significant at the .01 level for one-tail test

Table 4 summarizes the explanatory power as measured by the R-squared for each model. For the Summer 2005 course the R-squared for the proctored final is 61.6%, much higher than the R squared for the first three un-proctored exams, which average 30.5%. For the Summer 2004 course the R-squared for the proctored final is only 12.2%. These results suggest that the exam outcomes are much more correlated with human capital endowments for the proctored format than for the un-proctored format.

Table 4			
Summary of R^2 for OLS Results			
Exam	Full Sample	Summer 2004	Summer 2005
Exam 1	0.3560	0.4460	0.3483
Exam 2	0.2055	0.3547	0.1894
Exam 3	0.3687	0.0173	0.3781
Final Exam	0.3494	0.1219	0.6160

OAXACA DECOMPOSITION

The Oaxaca Decomposition (Oaxaca 1973) provides more information about the source of the differences between the final exam scores of the two courses. The Oaxaca Decomposition can be written as:

$$\text{Exam}_p - \text{Exam}_{up} = B_p\,(X_p - X_{up}) + (B_p - B_{up})\,X_{up}$$

where the subscript "p" is proctored, "up" is un-proctored, Exam is final exam score, B the vector of estimated coefficients, and X the vector of human capital endowments. The decomposition separates the proportion of the gap that is attributable to differences in human capital endowments $B_p\,(X_p - X_{up})$, and the proportion that is attributable to differences in returns to human capital endowments $(B_p - B_{up})\,X_{up}$.

The Oaxaca Decomposition for the difference between the mean final exam score for the Summer 2005 and Summer 2004 courses is calculated using the estimation results in columns 8 and 4 of Table 3, and the result is reported in Table 5.

Table 5		
Oaxaca Decomposition		
	Amount	Proportion
Difference in Final Exam Scores (2005 less 2004)	5.05	100%
Differences in Human Capital	0.87	17%
Differences in Returns to Human Capital	4.18	83%

Table 5 shows that only 17% of the gap is attributable to human capital endowments, and 83% is due to differences in returns to human capital. In other words, most of the difference in the final exam scores is explained by the higher returns to human capital in the proctored format, and very little of the gap is explained by the marginally higher human capital of the students in the course with the proctored format.

Suppose that the student that had enrolled in the Summer 2004 section had instead taken the Summer 2005 section with its proctored final exam, would his exam score be higher? That hypothetical is calculated by applying the estimated coefficients for the Summer 2005 sample to the sample data for 2004. The result is that the predicted mean test score would have been 3.51 points higher in the proctored format and the difference is statistically significant at the 5% level. We speculate that the cheating opportunities present in the un-proctored format encouraged students in the Summer 2004 course to study less than their counterparts in the Summer 2005 course causing their final exams scores to be lower.

CONCLUSION

This study addresses the question: Does mode of assessment format (proctored or un-proctored exams) affect test scores in online principles of economics classes? The data for the study data are from two courses of principles of macroeconomics, one taught in Summer 2004, the other in Summer 2005. The courses are identical in every respect, except the final exam in the Summer 2004 course was not proctored, and the final exam in the Summer 2005 course was proctored. Comparison of the mean final exam score of the raw data shows that the mean for the proctored final in the Summer 2005 course is 4.9 points higher than for the un-proctored final exam in the Summer 2004 course and the difference is statistically significant at the 10% level. A Oaxaca Decomposition is conducted with the result that 17% of the difference in test scores is due to differences in human capital endowments, and 83% is explained by differences in rates of return to the human capital endowments. It is speculated that the un-proctored exam format encourages students to study less and learn less than they would in the proctored format.

REFERENCES

Brown, B., W. and E. L. Carl (2002). "Can Web courses replace the classroom in principles of microeconomics?" The American Economic Review **92**(2): 444.

Carnevale, D. (1999). "How to Proctor From a Distance." Chronicle of Higher Education, **46**(12): A47.

Coates, D., Brad R. Humphreys, John Kane, and Michelle A. Vachris (2004). ""No Significant Distance" Between Face-to-Face and Online Instruction: Evidence from Principles of Economics." Economics of Education Review **23**: 533-46.

Coates, D. and Brad R. Humphreys (2001). "Evaluation of Computer-Assisted Instruction in Principles of Economics." Educational Technology & Society **4**(2): 444-9.

Deal, W. F., III (2002). "Distance learning: teaching technology online. (Resources In Technology)." The Technology Teacher **61**(8): 21(6).

Edlin, R., J. (2000). "Information technology in the classroom: experiences and recommendations." Campus - Wide Information Systems **17**(1): 10.

Kushner, D. (1999). Class Is in Session. Now What? New York Times**:** 4A.26.

Liefert, J. (2000). "Measurement and Testing in a Distance Learning COurse." Journalof Instruction Delivery Systems **14**(2): 13-16.

Navarro, P. (2000). "Economics in the Cyberclassroom." The Journal of Economic Perspectives **14**(2): 119-132.

Oaxaca, R. (1973). "Male-Female Wage DIfferentials in Urban Labor Markets." International Economic Review **14**(3): 693-709.

Rovai, A. P. (2001). "Online and traditional assessments: what is the difference? ." Online and traditional assessments: what is the difference? **3**(3): 141-151.

Serwatka, J. A. (2003). "Assessment in on-line CIS Courses." The Journal of Computer Information Systems **44**(1): 16.

Shuey, S. (2002). "Assessing Online Learning in Higher Education." Journal of Instruction Delivery Systems **16**(2): 13-18.

Sosin, K. (1997). "Impact of the Web on Economics Pedagogy." Presented at the Allied Social Sciences Association Meeting, January 5, 1997.

Taylor, S. S. (2002). "EDUCATION ONLINE: Off Course or On Track?" Community College Week **14**(20): 10-12.

Vachris, M. A. (1999). "Teaching Principles of Economics without "Chalk and Talk": The Experience of CNU Online." **30**(3): 292.

Young, J. R. (2001). "Texas Colleges Collaborate to Offer Online Students Convenient Proctored Tests." Chronicle of Higher Education, **47**(26): A43.