Journal of Business & Leadership: Research, Practice, and Teaching (2005-2012)

Volume 7
Number 1 *Journal of Business & Leadership*

Article 12

1-1-2011

On-Campus Vs. Online Course Delivery: An Empirical Look At Both Approaches In A Controlled Setting For Introductory Managerial Accounting

Win G. Jordan
Fort Hays State University

Amanda Brown
Fort Hays State University, albrown13@fhsu.edu

ON-CAMPUS VS. ONLINE COURSE DELIVERY: AN EMPIRICAL LOOK AT BOTH APPROACHES IN A CONTROLLED SETTING FOR INTRODUCTORY MANAGERIAL ACCOUNTING

Win G. Jordan, Fort Hays State University
Amanda Brown, Fort Hays State University

There continues to be much discussion about whether or not students learn as much in an online course as they do in an on-campus face-to-face setting. This paper presents empirical observations about four sections of introductory managerial accounting, two taught on-campus and two taught online. The on-campus face-to-face approach provided the same course content available in the online approach but also used classroom lectures and discussions. A comprehensive final exam covering all learning objectives of the course was used as the overall measure of content learning. The hypothesis was that content learning was not equal in the two groups, but a t¬-test using unequal variances indicated that essentially equal content learning was occurring under both approaches (the null hypothesis). Quantile regression also was used to uncover some insights not revealed by the t-test, indicating that among the worst-performing students, online students performed better than face-to-face students and that younger students outperformed older students.

INTRODUCTION

With the increase in availability of online courses, concern continues to exist on their effectiveness. Consider, for example, online courses at the authors' Midwest state comprehensive university. Twelve years ago, the counsel given to advisors and young students recommended against having young students take their courses online. At that time, the department charged with online education collected evidence showing students over 25 years of age generally outperformed younger students in online courses. Recently, however, that supposition does not appear to hold true. The question remains as to whether online students learn as much as on-campus face-to-face students. For the purposes of this paper, "learning" refers to the content embraced in the learning objectives; no claims are made about any non-content learning or other qualitative aspects that may occur in a course.

Many studies have been conducted on a variety of issues surrounding distance education vs. traditional classroom education. Latchman et al. (1999) started their paper by saying, "there is no doubt that nothing will replace synchronous learning with face to face interaction" (p. 247). However, as they conducted their study, they found that using the Internet to supplement classroom education could be an alternative to the classroom both for distance students and students unable to attend a particular class; enhanced learning occurred in either situation. Lectures, notes, and additional materials could be made available, much as occurs with Blackboard today.

In a report by The Institute for Higher Education Policy (Phipps & Merisotis, 1999), the authors stated:

> Most of these studies conclude that, regardless of the technology used, distance learning courses compare favorably with classroom-based instruction and enjoy high student satisfaction. For example, many experimental studies indicate that students participating in distance learning courses perform as well as their counterparts in a traditional classroom setting. These studies suggest that the distance learning students have similar grades or test scores, or have the same attitudes toward the course.

But the same study went on to recommend caution in placing too much assurance on that conclusion since the conclusion may not be generalizable across all courses.

Anstine and Skidmore (2005) examined whether MBA students who took only online courses learned as much as students taking identical courses in the traditional, face-to-face format. The MBA program that was studied offered economics classes, introductory statistics classes, and managerial economics classes in both online and traditional in-class formats. Comparison of test scores indicates similar content learning in the two formats. However, when they controlled for other non-content factors, use of a switching regression showed the online environment to be less effective than the traditional environment (p. 3)

Fajardo (2011) examined accounting courses – both online and face-to-face – with a focus on tying educational strategies and assessments to desired learning outcomes. Fajardo examined performance on the Standardized Learning Outcomes Assessment Tests (SLOATS) across several years. Each learning objective used four multiple-choice questions and one or two essays to assess performance. Each year the author examined each of the selected courses by learning method, considering the mean score of each method and the proportion of students in each method achieving at least 75% on the learning objectives. Not all courses were offered in both methods. In regards to online vs. onsite courses, Fajardo indicated that sometimes online would perform better, sometimes onsite would perform better, and sometimes they were about the same.

Neuhauser (2002) studied learning style and effectiveness of online and face-to-face instruction using two sections of a course. Neuhauser compared performance on each component of the course (test scores, assignments, participation grades, and final grades) but found no significant differences. Having focused on equivalency, Neuhauser turned to her primary interest: student perceptions of effectiveness of the current course's components in comparison with the effectiveness of courses taken under the other method. (Students taking a class online compared the effectiveness of each component with those of face-to-face courses taken. Students taking a class face to face compared the effectiveness of each component with those of online courses taken.) Neuhauser concluded that equivalent activities were equally effective in either situation. Thus the use of the course components focused on establishing course equivalency as a precondition to subsequent analysis.

Leasure et al. (2000) compared student outcomes from online and traditional sections of an undergraduate research course, finding no significant difference in examination scores between the two groups when using a t-test of exam score means. Of interest are the findings of Leasure et al. on the significant self-selection that occurred as students chose which approach to pursue. The reasons given for choosing the traditional classroom included perceptions of increased interaction, less opportunity to procrastinate, receiving immediate feedback, and more meaningful learning activities. Those choosing the online section cited cost, convenience, and flexibility. Leasure et al. found that students who were most suited to online courses were those reporting greater self-direction, the ability to pace themselves, and a lack of procrastination.

Vogel (2011) made an interesting twist by comparing performance of a course taught in three modes: online, in the traditional classroom, and as a hybrid course. Vogel evaluated a score production function (based on innate ability, student work input, and course mode) to assess outcome differences in the modes of presentation. In addition, Vogel also used ordinary least squares (OLS) and a Censored Tobit, finding that online students tended to perform better than traditional students, whereas traditional students performed slightly better than students in the hybrid class.

Johnson et al. (2000) compared outcome data from two sections, one online and one face-to-face, of a graduate-level instructional design course in the human resource development field. Both sections were taught by the same instructor with the same content, activities, and projects. Grade distributions in the two sections were essentially equal. Comparisons were based on the perceptions of students and included such items as student ratings of instructor and course quality and perceptions of interaction and support. Learning outcomes considered several other aspects in addition to exams. In all areas, no significant differences were found. The authors went on to infer that this lack of difference in learning outcomes supports the continued development and use of online programs.

Harrington (1999) compared performance of students in online and traditional statistics classes taken as part of a Master of Social Work program. Performance was based on homework assignments and closely related quizzes. No overall difference in performance was seen between classes in terms of being an online or traditional class. However, when Harrington subdivided the classes by student grade point average (GPA), a difference in performance was noted for lower-GPA students. Online students with higher GPA's did as well as higher-GPA students in the traditional class. Online students with lower GPA's did not do as well as lower-GPA students in the traditional class. The Harrington study contributed to the idea of looking for differences in segments of the population used in the current study with quantile regression.

After compiling dozens of studies on distance education, Russell (1999) found no difference in student learning. This appeared to be true especially in studies where different sections of the same class were taught both online and in the traditional classroom. Studies came from a number of disciplines. Of the studies discussed so far in the current article, Fajardo (2011) was in accounting, Leasure et al. (2000) was in research, Johnson et al. (2000) was in instruction design, and Harrington (1999) was in social work statistics. To these studies can be added Horiuchi et al. (2009) in continuing education for nursing and Thirunarayanan and Perez-Prado (2001) in ESOL (teaching English to speakers of other languages). In each of these cases, no significant overall difference was found between online and traditional face-to-face results. None of these studies indicated that they involved introductory managerial accounting.

Arbaugh (2000) said, "Comparing exam performance in asynchronous Internet-based courses with that of traditional classrooms will help determine whether the 'no significant difference' phenomenon prevalent with other forms of distance education can be extended to them as well" (p.216). Arbaugh went on to say, "Another significant limitation relates to the measures used in this study. A single measure of learning may not completely capture the content and quality of student learning experiences" (p. 227). Later, in Arbaugh (2005), the author states, "Because most comparison studies have focused on a single course or courses within a single discipline, it is impossible to make definitive assertions on the impact of disciplinary effect" (p. 58).

Thus, such comparisons continue on a case-by-case basis. The current study sought to extend the comparison of results into the arena of introductory managerial accounting. This study examined content learning of online vs. face-to-face approaches when taken as a whole to see if one provided superior results. The study also examined via quantile regression whether differences arise when considering segments or subgroups of the students.

https://scholars.fhsu.edu/jbl/vol7/iss1/12
DOI: 10.58809/VBVX4780

METHODOLOGY

The study covered four sections of the primary author's undergraduate Principles II (managerial accounting) class, two sections being on-campus face-to-face sections and two sections being online sections. The students self-selected into online or face-to-face sections as part of their normal enrollment process; thus the formation of the sections was not controlled. The on-campus sections contained a total of 50 students, whereas the online sections contained a total of 30 students. The on-campus sections were offered the same content available to the online students plus class lectures and discussions. All students in the class used Aplia as a third-party homework aid; exams were administered using the testing capabilities in Blackboard. A comprehensive multiple-choice final covering all learning objectives was used as the overall measure of content learning. Because of the inability to control the testing environment of online students, all students took the exam at home, maintaining the same conditions for both on-campus and online students. Similarly, the exams for all sections were open-book and open-note with a time limit of 75 minutes. The current study used the same rationale as that expressed by Anstine and Skidmore (2005): "Because online students, by the nature of the learning environment, can use resources such as textbooks and notes to answer test questions, the students in the traditional class were also given take-home exams and were allowed to use authorized resources to answer test questions" (p. 112). To further control the testing environment in the current study, questions on the exam were randomly pulled from a pool for each chapter covered.

It should be noted that all students worked under the same examination set up, thus allowing for a better comparison of final exam scores between the two groups. This equality in testing conditions was not present in Arbaugh (2005), Fajardo (2011), Johnson et al. (2000), or most of the other studies discussed above because they were focusing on different matters. Other authors, such as Neuhauser (2002), did not address the testing conditions. Harrington (1999) used the same quiz questions for both methods, but no mention was made of using a final, especially a comprehensive one. In Arbaugh (2000), student learning was measured using a 50-question multiple choice exam. Anstine and Skidmore (2005) used identical exams for a particular course, but different types of exams for different courses, with all being take-home exams.

A simple t-test was performed on the mean final exam scores of on-campus students and online students. The authors put forth the hypothesis that there would be a difference between the means, thus the hypotheses became:

H1: A difference exists between the means of online student final scores and on-campus student final scores.

H0: No difference exists between the means of final scores for the two groups.

The authors considered whether students' self-selection had resulted in atypical groups. Therefore the authors attempted to determine possible differences in the student population that would account for the results. In data collection, the following demographic attributes had been requested from students: gender, age, major, entering cumulative GPA, whether they were primarily an on-campus student or an online student, credit hours taken that semester, weekly work hours, and country of origin The demographic attributes were examined to see whether correlations existed between them and performance on the final exam. (Attributes with strong correlation could be examined to determine if the two populations differed.) In this process, multiple regression and quantile regression were used.

Regression analysis looked at the population as a whole or as if involving a "typical" student. Multiple regression coefficients indicated the average change in the dependent variable given a change in a specified independent variable, but they failed to consider potential differences between subgroups of the population. All members of the population were assumed to be functioning alike. Regression analysis was used primarily to identify which variables significantly contributed to the final exam score. The independent predictor variables included the following:

1) Aplia points, a continuous variable, reflected accomplishment on homework and was expected to contribute positively to the final exam score.

2) Total credit hours entering the course was a category variable indicating the student's class standing and was expected to be negatively correlated since higher class standing (junior or senior) indicated that the student had postponed taking the course instead of taking it when expected (usually sophomore year).

3) On-campus versus online was a dummy variable (on-campus = 1; online = 0) that captured the learning format in which the student was taking the class; a positive correlation was expected originally.

4) Gender was a dummy variable (male = 1; female = 0) expected to have no significant correlation to final exam scores.

5) U.S. versus foreign was a dummy variable (U.S. = 1; foreign = 0) expected to have no significant influence.

6) Average exam score was a continuous variable expected to correlate positively with final exam scores.

7) Working 5 to 20 hours per week was a dummy variable (yes = 1; no = 0) expected to have a negative correlation with final exam scores.

8) Working over 20 hours per week was a dummy variable (yes = 1; no = 0) expected to have an even stronger negative correlation with final exam scores since it reflected having less time to study.

9) Age was treated as a category variable (over 30 = 5; 26-30 = 4; 23-25 = 3; 20-22 = 2; under 19 =1) which we originally expected to be positively correlated with final exam scores.

10) Overall GPA entering the course was treated as a category variable (under 2.0 = 0; each additional half point of GPA went up 1 until the highest category was 3.5-4.0 = 4) expected to have a positive correlation with final exam scores.

11) More than 12 credit hours this semester was a dummy variable (yes = 1; no = 0) expected to have a positive correlation since it indicated a full-time student who might have less time demands than a part-time student.

Consideration was given to whether variations within the population might yield additional insights. Harrington (1999) showed an example where results were better interpreted by considering portions of the population. After finding no difference between online and traditional students as a whole, Harrington subdivided the classes by student grade point average (GPA). Examining the students in this new context, a difference in performance was noted; online students with lower GPA's did not do as well as traditional lower-GPA students.

In the current study, quantile regression was used to examine different performance in different subgroups. By subdividing the population into levels of the dependent variable (final exam score in this study), the various levels of performance were examined to determine where, if any, particular independent variables had a significant effect on that subgroup. This refinement provided additional insight into the relationship between a variable and resulting performance. Thus high-performing students were contrasted with middle-performing students and low-performing students to see if a variable impacted one subgroup more than another.

RESULTS

As shown in Table 1, using a t-test of two means with unequal sample sizes and unequal variances found no significant difference. Tests showed that assuming unequal variance was justified. However, examination assuming equal variance yielded nearly identical results (with the resulting probability being .20 instead of .19). Therefore, the H1 hypothesis was rejected; no significant difference was found in the content performance of on-campus students in comparison with that of online students. A simple t-test, however, does not control for other factors that potentially might explain any differences in the two groups. This led to performing a regression analysis, as discussed above.

TABLE 1

On-Campus vs. Online Final Exam Scores t-Test: Two-Sample Assuming Unequal Variances

	On-Campus	Online
Mean	65.54	69.93
Variance	242.25	188.54
Observations	50	30
Hypothesized Mean Difference	0	
df	67	
t Stat	-1.31688	
P(T<=t) two-tail	0.192364	
t Critical two-tail	1.996008	

https://scholars.fhsu.edu/jbl/vol7/iss1/12
DOI: 10.58809/VBVX4780

When the OLS regression analysis was run including all the variables, only four variables and the constant showed any significant relationship to final exam scores. As shown in Table 2, a strong positive relationship (at the .001 level) existed between average exam scores and final exam grades; this was as expected. A significant negative relationship (at the .01 level) was found between on-campus students and final exam scores, indicating that online students might be performing better. A weak positive relationship (at the .1 level) existed for total credit hours taken that semester, as was expected. In addition, a weak negative relationship (at the .1 level) was noted for weekly work hours, as was expected. Note that the lack of significance between age and final exam score supported the belief that age was no longer an important factor in determining whether students should be advised to take online courses.

TABLE 2

Results of Multiple Regression to Identify Variables Influencing Final Exam Scores

Variable	Coefficient	1-tailed p
(Constant)	26.306	*
Aplia points	.139	
Total credit hours entering course	-.045	
On-campus vs. online (dummy variable)	-9.823	**
Gender (dummy variable)	1.299	
U.S. vs. foreign (dummy variable)	-5.077	
Average exam score	.723	***
Working 5 to 20 hours per week (dummy variable)	-1.616	
Working over 20 hours per week (dummy variable)	-3.248	*
Age	-.299	
Overall GPA entering course	-.477	
More than 12 credit hours this semester	4.548	*

In order to work with quantile regression, the authors started with all the variables listed above, but the small sample size necessitated first identifying attributes (variables) with the greatest significance. Elimination of attributes found not to be significant led to the three variables and the constant: the constant (C), VIRTUAL_1, AGE, and EXAM_AVE. One change was made on how VIRTUAL_1was valued in the quantile regression; if the student was online, the value was 1, whereas a traditional student used a value of 0. AGE continued to use the categories described previously, and the average of the previous four exams in the class was shown as (EXAM_AVE). As shown in Table 3, these four attributes were found to be have overall significance in the least squares portion of the quantile regression. Since the constant is not of interest, it will not be discussed further.

TABLE 3

Least Squares Regression Results and Quantile Regression Results

	OLS Regression				Quantile Regression									
Variable	Coeff.	Std. Error	t-Statistic	Prob.	Coeff./ Signif. (*)	10th	20th	30th	40th	50th	60th	70th	80th	90th
C	20.390	8.475	2.406	0.019	C	-1.301	0.551	14.253	17.444	24.365	29.690	30.272	39.701	57.268
					S							*	*	*
VIRTUAL_1	5.290	2.946	1.794	0.077	C	7.952	6.933	6.997	4.667	7.173	6.675	3.952	0.546	-0.703
					S	*	*	*						
AGE	(.389)	0.208	(1.870)	0.065	C	-0.301	-.495	-0.562	-0.556	-0.250	-0.304	-0.192	-0.163	-0.152
					S	*	*	*	*					
EXAM_AVE	0.834	0.112	7.445	0.000	C	0.910	1.031	0.889	0.889	0.731	0.693	0.704	0.619	0.446
					S	***	***	***	***	***	***	***	***	**

R-squared	0.444
Adjusted R-squared	0.422
Log likelihood	-306.018
F-statistic	20.219
Prob (F-statistic)	0.000
Schwarz criterion	7.870

The quantile regression stratified the population into ten segments for each attribute of interest and examined its significance as a predictor of final exam scores. The coefficient showed the degree of change on final exam scores associated with a unit change in the attribute. The rate of change of the coefficient showed whether change occurs more rapidly in higher or lower segments.

Thus, considering VIRTUAL_1, shows that the impact on being an online student had a higher impact if the student was in the lowest-performing 30% of students. For those 30% of the students, online students actually did better than face-to-face students. One possible explanation is that the online students felt the pressure to study the materials more than their face-to-face counterparts since they had less instructor interaction. Whether or not the student was on-campus or online had no significant impact on the upper 70% of students.

Whereas the least squares regression showed AGE to be significant, the quantile regression revealed that the significance was confined to the worst-performing 40% of students in the population. For those 40% of the students, the negative correlation shows that younger students outperformed older students. One possible explanation is that the older students were taking on more responsibilities, such as employment, marriage, and family.

The EXAM_AVE variable was found to be highly significant in all segments of the population. It is not surprising that performance on previous exams would serve as an indicator of student performance on the final exam. However, the rate of change of the coefficient is decreasing at the higher segments of the population. This indicates that a single point on one of the earlier exams is less important in predicting the final exam score for a top student than it would be for someone else.

CONCLUSIONS

For the course examined (introductory managerial accounting), there was no significant difference in the final exam scores of on-campus vs. online students. The decision of students to take online courses as opposed to on-campus face-to-face courses may reflect a degree of self-selection, but further research would be needed to identify the attributes involved. Perhaps such work might consider Anstine and Skidmore (2005), where three variables were identified that "might influence a student's learning environment choice but not necessarily affect class performance": (a) travel time to a university, (b) whether the student had children in the home, and (c) reported weekly hours devoted to work. All three of the suggested variables are fairly common in students at the current authors' university, whether they are taking classes online or in the traditional classroom. Another possible variable would be how long it had been since taking college algebra and how well the student scored in that course. Gathering the data would permit further examination as to their impact on self-selection.

It must be stressed that this study was limited to examining the learning of content. No claims are made about any non-content learning or other qualitative aspects that may occur in a course. In Arbaugh (2000) the author warns, "Another significant limitation relates to the measures used in this study. A single measure of learning may not completely capture the content and quality of student learning experiences" (p.227). The same limitation applies to the current study. Finding the appropriate measures to use could offer opportunities for additional research.

Contributions of this study are two-fold. First, the control of conditions for students taking the final exam was identical, allowing for better comparison of the scores and their indication of content learning. This control often was missing in other studies because they were focusing on

https://scholars.fhsu.edu/jbl/vol7/iss1/12
DOI: 10.58809/VBVX4780

different matters. Second, the current study used quantile regression to provide additional insights by segmenting the population by performance level and looking at the impact of the variables on performance within each segment. In contrast, regular OLS regression simply considered the average impact of a variable on performance of the population as a whole.

REFERENCES

Anstine, J., & Skidmore, M. (2005). A small sample study of traditional and online courses with sample selection adjustment. *Journal of Economic Education*, *36*(2), 107-127.

Arbaugh, J. B. (2000). Virtual classroom versus physical classroom: An exploratory study of class discussion patterns and student learning in an asynchronous internet-based MBA course. *Journal of Management Education, 24*(1), 213-233.

Arbaugh, J. B. (2005). How much does "subject matter" matter? A study of disciplinary effects in on-line MBA courses. *Academy of Management, 4*(1), 57-73.

Fajardo, C. (2011). Evaluation of learning outcomes in undergraduate onsite and online accounting courses. *The Journal of American Academy of Business, Cambridge*, *17*(1), 18-24.

Harrington, D. (1999). Teaching statistics: A comparison of traditional classroom and programmed instruction/distance learning approaches. *Journal of Social Work Education 35*(3), 343–52.

Horiuchi, S., Yaju, Y., Koyo, M., Sakyo, Y., & Nakayama, K. (2009). Evaluation of a web-based graduate continuing nursing education program in Japan: A randomized controlled trial. *Nurse Education Today, 29*(2), 140–149.

Johnson, S. D., Aragon, S. R., Shaik, N., & Palma-Rivas, N. (2000). Comparative analysis of learner satisfaction and learning outcomes in online and face-to-face learning environments. *Journal of Interactive Learning Research*, *11*(1), 29-49.

Latchman, H., Salzmann, C., Gillet, D., & Bouzekri, H. (1999). Information technology enhanced learning in distance and conventional education. *IEEE Transactions on Education*, *42*(4), 247-254.

Leasure, A. R., Davis, L., & Thievon, S. L. (2000). Comparison of student outcomes and preferences in a traditional vs. world wide web-based baccalaureate nursing research course. *Journal of Nursing Education, 39*(4), 149-154.

Neuhauser, C. (2002). Learning style and effectiveness of online and face-to-face instruction. *American Journal of Distance Education*, *16*(2), 99-113.

Phipps, R., & Merisotis, J. (1999). What's the difference? A review of contemporary research on the effectiveness of distance learning in higher education. Washington, DC: www.ihep-.com/PUB.htm

Russell, T. (1999). *The No Significant Difference Phenomenon*. Raleigh, NC: IDEC.

Thirunarayanan, M.O., & Perez-Prado, A. (2001). Comparing web-based and classroom-based learning: A quantitative study. *Journal of Research on Technology in Education*, *34*(2), 131-137.

Win G. Jordan is an Assistant Professor of Accounting at Fort Hays State University in Hays, Kansas. After receiving his MBA at Brigham Young University, he received his Ph.D. at Michigan State University. Interspersed with his education, he has 13 years in industry, where he became intrigued with how decisions are made and how accounting information influences those decisions. His areas of interest are management accounting and accounting information systems, with special emphasis on how those two areas interact. He may be contacted at wjordan@fhsu.edu.

Amanda Brown is a graduate research assistant (GRA) in the MBA Program at Fort Hays State University in Hays, Kansas. After receiving her B.A. in psychology at Fort Hays State University, she entered the MBA Program at the same school. She finds ample opportunity to use her research and statistical skills as a GRA.