**The Effect of Online Learning on Student Performance:**

**A Replication Analysis of: *A Randomized Assessment of Online Learning (Alpert et al., 2016)***

Data Analysis Project Proposal

Gian Zlupko

HUDM 5123: Linear Models and Experimental Designs

Dr. Bryan Keller

November 21, 2021

**Context**

The data presented in this proposal were originally published in *A* *Randomized Assessment of Online Learning* (Alpert et al., 2016). In the original study, the researchers conducted a randomized study that examined the effect of instructional format on learning outcomes of college students. In particular, Alpert et al. (2016) examined the effect of three different learning formats (e.g. face-to-face, hybrid, and purely online) on students’ cumulative final grade in a university-level economics class.The researchers found that students’ final grades were significantly different across traditional, hybrid, and purely online learning formats. Specifically, students in the online learning condition received final grades that were, on average, four to five points lower than students in the traditional, in-person lecture format after controlling for GPA, SAT scores, and student credit hours.

Alpert et al. (2016)’s study is part of a substantial body of educational research that has examined the impact of online learning on student performance (Bowen et al., 2014; Figlio et al., 2013). Yet, despite the wealth of research on the effects of online learning, Alpert et al. (2016)’s study is among only a handful of randomized experiments in this literature. Moreover, their findings are particularly relevant today as universities are considering the costs and benefits of increasing online learning offerings for students, especially during the COVID-19 pandemic. Therefore, in light of the timely importance of Alpert et al. (2016)’s study, this paper proposes to conduct a replication study to evaluate the researchers’ findings. Specifically, I will examine the effect of instructional format on the final course grades that students receive. In addition, GPA will be incorporated as a covariate in the relationship between instructional format and final grades and, as discussed in greater detail below, multiple comparisons will be conducted to examine differences between instructional formats on final grades.

**Citation**

Data from the Alpert et al. (2016) study was publicly available and accessed online at Open ICPSR.org. A full citation for the original research study is as follows:

Alpert, B. W. T., Couch, K. A., & Harmon, O. R. (2016). A Randomized Assessment of Online Learning. *American Economic Association*, 106(5), 378–382.

**Variables**

The randomized treatment variable in this study was instructional format. Students interested in taking a college-level economics were randomly assigned one of three instructional format conditions; face-to-face classroom instruction, hybrid, and online. Students assigned to the face-to-face instructional format received traditional classroom instruction. For example, in in this condition, students attended two weekly lectures in-person for the entirety of the semester. In the hybrid format, students met in-person for one lecture per week and then were given online lecture materials to in place of a second in-person lecture. Finally, students assigned to the online format were provided the same online lecture materials as the hybrid section but were not given an in-person lecture. For each of the three conditions, lectures and instructional content were prepared and delivered by the same instructor.

The dependent variable in the Alpert et al. (2016) study the cumulative final grade that students received in the course. Cumulative grades were measured continuously out of 100 points total. In addition to analyzing the relationship between instructional format (IV) and cumulative grade (DV), the researchers also measured GPA and SAT as covariates. GPA and SAT were both measured continuously. For SAT scores, Alpert et al. (2016) controlled for both verbal reasoning and math SAT scores. Finally, the researchers also controlled for student demographics (e.g. race, gender) as well as cumulative university credits that students had acquired prior to enrolling in the course.

**Missing Data**

Alpert et al. (2016) randomly assigned students to instructional format conditions prior to the beginning of the semester. This meant that, naturally, not all students who were given permission to enroll in the class decided to enroll. Moreover, not all of the students who enrolled in the course completed the course. Therefore, there are two measures of attrition that are captured in the data set; ‘not enrolled’ and ‘dropped course’. Descriptive statistics for measures of attrition, “Did Not Enroll” and “Dropped Course”, are presented in Table 1 along with the total group sample size randomly assigned to each condition as well as the final group sample size at the end of the study.

**Table 1**

*Attrition by Group*

Table

Description automatically generated

The researchers chose following recommendations from similar randomized experiments on online learning, hence the decision to randomly assign participants to condition prior to the start of the school semester (Bowen et al., 2014; Figlio et al., 2013; Joyce et al., 2015). Doing so enabled the researchers to also control for and examine whether the mode of instructional format offered to students had an effect on nonenrollment and course drop out. Students that were randomly assigned to the option of taking the class online dropped the course at higher rates than students in the hybrid and traditional lecture formats. Alpert et al. (2016) used this attrition data in the original study to control for the possible confound of the effect of differential attrition. When controlling for attrition, the researchers found the same similar group mean estimates for final grade received and the same statistical significance obtained in an ANCOVA test without an attrition term included.

In addition to final grades, there are also missing records in the data set for student GPA. Because this proposal outlines the use of GPA as a covariate in the analysis of the effect of instructional format on students’ final course grades, careful attention will be paid to the missing GPA records in the data set. A review of missing GPA data by condition and the missing rate by group sample size are presented in Table 2.

**Table 2**

Table

Description automatically generated

Overall, 9 GPA records were missing in the online condition, 13 were missing in the hybrid condition, and 26 were missing in the traditional condition. Follow up steps in the replication study proposed herein will be taken to evaluate whether GPA data was missing at random (MAR). To test patterns of missingness, relevant functions from the R packages sjmisc, mi, DescTools, and VIM, will be used to evaluate whether there are patterns associated with GPA missingness. Then, based on these findings, an appropriate strategy for handling missing records will be implemented.

**Univariate Plots and Descriptive Statistics**

**Figure 1**

*Boxplot of Final Course Grade by Group*

**![Chart, box and whisker chart

Description automatically generated]()**

**Figure 2**

*Histograms of Final Course Grade by Group*

![Chart, histogram

Description automatically generated]()

**Figure 3**

*Student GPA Distribution by Group*

**Chart

Description automatically generated**

**Figure 4**

*QQ-Plot of Normality for Final Grades*

**![Chart, line chart

Description automatically generated]()**

**Table 4**

*Final Grade Descriptive Statistics*

*Table

Description automatically generated*

**Table 5**

*Student GPA by Group Descriptive Statistics*

*Table

Description automatically generated*

**Research Questions**

RQ1: The purpose of this project will be to examine the effect of instructional format on student grades. Specifically, I will test whether the type of instructional format that a student received had an effect on their final grade. In addition, since a student’s GPA and course grade are expected to be highly correlated, I will also control for GPA as a covariate in the relationship between instructional format and student grade.

RQ2: The second purpose of this study will be to follow up the initial analysis discussed above by conducting a series of comparisons to examine pairwise differences between instructional formats on final grades. In doing so, my goal is to identify which instructional formats are significantly different from one another with regards to student learning outcomes. Moreover, for those comparisons that are significantly different, I will estimate the predicted difference in students’ final grade by instructional format.

**Statistical Methods**

To address RQ1, I will conduct one-way ANCOVA to estimate the treatment effect of instructional format on students’ final grades while controlling for GPA. Specifically, I will conduct an omnibus test of a one-ANCOVA. To address RQ2, I will conduct follow up pairwise comparisons of group means to determine whether certain instructional formats are significantly different from one another. To control for familywise Type I error rate at α = .05, I will conduct Tukey’s HSD adjusted pairwise comparisons.

**Assumptions**

In order to conduct ANCOVA, I will need to satisfy the general assumptions required for ANOVA (εi iid∼ N(0,σ), and I will also need to satisfy ANCOVA-specific assumptions. First, to test ANOVA assumptions, I will need to i) demonstrate that residuals are normally distributed (Normality), ii) demonstrate that variance is constant across the study groups (Constant Variance), iii) demonstrate that residuals have a mean of zero across all values of focal variables (Linearity), and iv) demonstrate that residuals are mutually independent. To demonstrate that these conditions are satisfied, I will generate boxplots and histograms to visually inspect whether variance is constant across groups and to determine whether the data is normally distributed. In addition, I will also generate a QQ plot to examine whether student grades were normally distributed. Also, given unequal-n in this study, I will use Maxwell et al. (2017)’s rule of thumb for assessing constant variance. Specifically, the rule states that constant variance can be established if the product of the ratio of the largest subgroup to the smallest and the ratio of the largest variance to the smallest is less than four. Finally, for a statistical test of constant variance, I will also calculate Levene’s test for homogeneity of variances. The assumption of independence has been satisfied through the random assignment of participants to groups.

In addition to checking for ANOVA assumptions, I will also need to satisfy two ANCOVA-specific assumptions, a) homogeneity of regression slope, and b) that the assumption that the covariate and dependent variable are linearly related. To test for homogeneity of regression slopes, I will generate a scatter plot and fit the regression line estimated by GPA for each group. The scatter plot with fitted lines will then be visually inspected to determine whether the regression lines are parallel. I will follow up this visual inspection of homogeneity of regression slopes by conducting a statistical analysis of regression heterogeneity. Specifically, I will estimate an interaction term in the ANCOVA model and inspect whether the interaction term is significant. The term should not be significant in order to demonstrate heterogeneity of regression slopes. Finally, to test for a linear relationship between the covariate and the outcome variable, I will visually inspect the aforementioned scatterplot to evaluate whether the covariate and outcome are linearly related at every level of the outcome variable across each of the three study groups. In addition, I will calculate the correlation between the covariate and the outcome to demonstrate that the relationship between the covariate and the outcome is large enough (*r* ≥ .30 ) to account for substantial variance explained in the ANCOVA model, demonstrating suitability for a covariate.

References

Alpert, B. W. T., Couch, K. A., & Harmon, O. R. (2016). A Randomized Assessment of Online Learning Author. *American Economic Association*, *106*(5), 378–382.

Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2014). Interactive learning online at public universities: Evidence from a six-campus randomized trial. *Journal of Policy Analysis and Management*, *33*(1). https://doi.org/10.1002/pam.21728

Figlio, D., Rush, M., & Yin, L. (2013). Is it live or is it internet? experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics*, *31*(4). https://doi.org/10.1086/669930

Joyce, T., Crockett, S., Jaeger, D. A., Altindag, O., & O’Connell, S. D. (2015). Does classroom time matter? *Economics of Education Review*, *46*. https://doi.org/10.1016/j.econedurev.2015.02.007

Maxwell, S. E., Delaney, H. D., & Kelley, K. (2017). Designing Experiments and Analyzing Data. In *Designing Experiments and Analyzing Data*. https://doi.org/10.4324/9781315642956

.. accessed from Open ICPSR here: [A Randomized Assessment of Online Learning](https://www.openicpsr.org/openicpsr/project/113462/version/V1/view)

**Appendix A**

Online Resources

R code, data visualizations, and a data set for this replication study can be accessed via the publicly-available GitHub repository below:

<https://github.com/gzlupko/5123_Linear_Models/tree/main/Online_Learning_Study>

The data from Alpert et al. (2016)’s A Randomized Assessment of Online Learning was originally accessed online at Open ICPSR at the following URL:

<https://www.openicpsr.org/openicpsr/project/113462/version/V1/view>