The Interaction of OER Use and Course Difficulty on Student Course Grades in a Community College

Introduction

A community college in Virginia, USA, has developed a ZTC degree in which it is possible to complete all coursework for the degree with zero textbook costs. The term ZTC simply emerged from how sections of courses are listed in the course schedule. Some sections of courses require a commercial textbook and some sections of the same courses utilize OER. Sections that use OER are labeled in the schedule with a lowercase “z” beside the section number. Because many courses have multiple sections - some which require either commercial textbooks and some which use OER, it is possible to analyze potential differences in outcomes controlling for student attributes and estimating interaction effects with course attributes such as course difficulty. This study was conducted to test such course outcomes and interactions.

Review of Literature

Most college instructors require students in their courses to obtain learning materials (Seaman & Seaman, 2017), and the price of commercial learning materials, particularly textbooks, has increased dramatically in the past few decades (US Bureau of Labor Statistics, 2016). An alternative to expensive commercial materials are Open Educational Resources (OER), which include a variety of available learning materials such as textbooks, music, and videos that are licensed without access fees (Butcher, 2015) and are openly licensed for retention, reuse, revision, remixing and redistribution.

The COUP framework (i.e., Cost, Outcomes, Usage, and Perceptions) has been used to evaluate OER (Bliss et al., 2013). Beyond estimates of costs and savings (C), usage (U), and perception
(P), a critical aspect to consider are the outcomes (O). If students save money, usage is widespread and nuanced, and perception is favorable, but student learning is not on par with the use of traditional textbooks, then the benefit of OER is diminished.

Most studies of OER outcomes have shown that courses using OER have comparable learning outcomes with courses using traditional textbooks (e.g., Clinton & Kahn, 2019). Sometimes the outcomes for OER are better and occasionally they are worse. Reviews by Hilton (2016, 2019) concluded that students generally achieved the same learning outcomes in classes with OER, compared with students in classes with non-OER. Robinson (2015) utilized a quasi-experimental design to compare student learning outcomes between sections in the treatment group (OER) and sections in the control group (non-OER) among seven different courses. Overall, five sections using OER showed similar or better outcomes than sections of the same courses using traditional textbooks. Two sections of courses showed better outcomes using traditional textbooks. The same mixed pattern can be also found in a multi-institutional study by Fischer et al. (2015). The authors utilized propensity score matching to control for age, gender, and minority status in 15 courses. Each course had sections that used either a traditional textbook or OER. The majority of courses (10) showed no difference in student grades according to OER vs. traditional textbook used. Four courses showed better grades in OER sections and one course showed better grades in the section using the traditional textbook.

A meta-analysis that aggregated findings from 22 studies with a combined total of over 100,000 students in which OER textbooks were compared to traditional textbooks found that learning outcomes were equivalent (Clinton & Kahn, 2019). However, there was substantial variability across studies in effect sizes of learning outcomes between OER vs. non-OER. All of the studies used quasi-experimental designs with varying levels of control for possible confounds, such as being taught by different instructors. The authors grouped the studies for three potential methodological confounds: whether or not there was the same instructor, whether or not the same learning measurement was used to measure outcomes, and whether or not prior knowledge or academic achievement was accounted for in the findings. The findings on learning outcomes did not vary based on whether those potential confounds were accounted for. Therefore, it is uncertain why there was so much variability in learning outcomes across studies. However, when considering the relatively small effect sizes attributed to textbooks in general (Robinson 2015) and the typically low coefficients of determination, it becomes apparent that variability in student performance is associated with myriad unmeasured covariates.

The access hypothesis provides a useful understanding of the meta-analytic findings on open textbook adoption. According to the access hypothesis, having access to learning materials
would be advantageous to learning outcomes; however, the number of students who would not have access to commercial resources but whose learning would benefit from access is relatively small (Grimaldi et al., 2019). Therefore, the effect of OER adoption on learning outcomes averaged across all students in all courses is likely to be null, as was found in the meta-analysis by Clinton and Kahn (2019). However, Grimaldi and colleagues (2019) commented that it is important to consider how different contexts may vary the outcomes of OER adoption, which is also evident by the large variability in effect sizes in Clinton and Kahn (2019).

One area in which the context interacted with OER adoption on learning outcomes was with student socioeconomic status. Two studies on OER adoption found that students who were eligible for a certain type of financial aid based on low-income status (Pell grants) benefited from OER adoption more than their peers (Colvard et al., 2018; Delgado et al., 2019). This is consistent with the articulation of the access hypothesis by Grimaldi et al. (2019) because students who had less income likely had fewer financial resources for course materials than their peers and may have been less likely to access pricey commercial resources, but could access the OER available without fees. Their peers may have been able to afford the commercial materials and received less benefit from OER adoption because they were able to access both commercial resources and OER.

There has been some examination of different contexts for outcomes of OER adoption. No extant study has examined how course difficulty may relate to OER and student learning outcomes. Approximately one-third of students in a study reported that not having the textbook due to cost had negative academic consequences (Florida Virtual Campus, 2018). Perhaps the use of OER in more “difficult” courses has a differential effect on outcomes because the potential effects of not having a textbook would be greater with more challenging courses. Granted, what is difficult for one student might be quite easy for another. Rather than stereotype departments and courses as difficult or easy, we acknowledge the fit between student interest and talents and the courses they complete. Nevertheless, some reasonable estimate of course difficulty might be important to consider in estimating the outcomes associated with the presence of OER.

Researchers have tried various approaches to estimate course difficulty but have mostly relied on perceptions of students or researchers. Ridley et al. (2003) used the perceived severity of grading standards to estimate intellectual challenge and course difficulty. Similarly, Bassiri et al. (2003) used grading policy in syllabi to estimate course difficulty. Babad et al. (2008) estimated course difficulty by analyzing perceived workload from course syllabi. Interestingly, Ansburg (2001) used student expectations of grade distributions to estimate course difficulty, where the logic was that a course that was of appropriate difficulty would have a negatively skewed distribution of grades. They expected that grades would generally be
on the high end with few low grades in the class. The students’ expectation was that more difficult courses would have a normal distribution around a mean of 2.0 with fewer A grades. The idea of using distributions of grades seemed to be a reasonable approach to quantitatively estimate course difficulty. Indeed, Anderson et al. (2018) estimated course difficulty using historical grades and withdrawal rate in two finance courses (two sections each). While the withdrawal rate did not accurately discriminate between the two courses, the historical grade distributions seemed to be an appropriate discriminator. Wladis and Hackey (2014) estimated course difficulty simply by distinguishing between “lower level” courses and “higher level” courses based on the presence of credit-bearing pre-requisites. If a 200-level course had a credit-bearing pre-requisite, it was deemed to have higher difficulty. The authors did not find a significant effect of online versus face-to-face delivery on retention rates in higher level courses.

In addition to examining how OER outcomes may vary depending on context, another area in need of development is controlling for confounding variables. Because of the pragmatic realities of conducting research with college courses, quasi-experiments comparing naturally occurring groups (students enrolled in different courses) are typically the methodology used. This methodological approach allows for ecologically valid comparisons because real students in real courses are examined. However, the lack of random assignment in quasi-experiments limits the likelihood the compared groups were similar in important characteristics such as demographics or prior academic achievement. For these reasons, Clinton’s review of OER in psychology courses (2019) called for better control of potential confounds as this lack of control is a valid critique of OER efficacy research (see Griggs & Jackson, 2017; Gurung, 2017). Indeed, Clinton (2018) found that differences in prior academic achievement likely explained differences in learning outcomes when comparing an introduction to psychology course with a traditional textbook to one with an OER textbook. Some studies have controlled for possible confounds. For example, Fischer et al. (2015) used propensity score matching to control for age, gender, and minority status across all courses. In addition, Jhangiani et al. (2018) measured prior knowledge preceding the study and found that students in different courses had comparable background knowledge.

The current study was a test of the interaction between OER and course difficulty in a robust sample of courses and students while controlling for potential confounds. The primary research questions were:

1. **What is the association of textbook type with students’ course grades controlling for gender (self-reported), Pell grant eligibility (as a proxy for student socioeconomic status, see Colvard et al., 2018, for a similar approach), prior academic success, and course difficulty?**

2. **Does the association of textbook type
with students’ course grades vary with course difficulty?

Prior academic performance is particularly important to control for because it is such a strong predictor of performance on learning assessments (Cassidy, 2015).

Method

The study was conducted in a community college in Virginia that has adopted an OER-based pedagogy that allows students to earn associate degrees with zero dollars spent on textbooks (DeMarte & Williams, 2015; Wiley, Williams, DeMarte, & Hilton, 2016). Data were obtained from 35 courses, which had both non-OER and OER sections, offered during the summer and fall semesters of 2016. Those courses were taught by 388 instructors. Some of the instructors taught courses or sections in the ZTC degree with OER and also taught courses outside of the ZTC degree with traditional textbooks. The courses included a wide range of subjects including business, mathematics, computer programming, biology, chemistry, history, music, and sports, which was a representative list of courses offered in a community college. Approximately 25,117 course grades were included but with listwise deletion of data based on the eventual covariates considered, 15,633 course grades were considered. Data were extracted from the college’s archives.

The dependent variable, Course Grades, estimated students’ learning outcomes and were reported on a five-point scale, A, B, C, D, and F (4,3,2,1,0). Five independent variables were included in the study: OER Course (Yes/No), Gender (Male/Female), Pell Eligibility (Yes/No), Course Difficulty (continuous) and Previous GPA (continuous).

OER Course was measured as a binary variable with 1 being OER course and 0 being non-OER course. Self-reported gender in the system was binary, male and female. Pell eligibility (1: eligible; 0: not eligible) and prior GPA were extracted for each student from the college’s records. Prior GPA was standardized to a z-score, which has a mean of 0 and standard deviation (SD) of 1 (original mean = 2.94; SD = 0.78). The course difficulty variable was based on failure rates in the current courses. It was created by calculating the proportion of students achieving a D grade or lower across all sections of each course (e.g., if 80% of students who took the course received a D or lower grades, the difficulty would be 0.8). Course difficulty was then standardized (i.e., standardized difficulty = (raw difficulty – mean difficulty of all courses) / SD of all courses) around the mean failure rate of 0.28 (SD = 0.8; Range, 0.08 to 0.43) to render a continuous variable with mean of 0 and SD of 1. Hence, the larger the difficulty score, the more difficult the course was, and positive course difficulty scores (i.e., above mean) meant that the course was more difficult than the courses with negative difficulty scores (i.e., below mean).

The purpose of standardizing the two continuous variables (prior GPA and course difficulty) was for inter-
pretability of results. Standardizing the two continuous variables created an interpretable zero-point. The remaining three variables OER use, Pell eligibility, and Male were binary and coded with an interpretable zero. Standardizing the two continuous independent variables made interpretation more consistent with the interpretation of binary variables, that is, the estimated change in the outcome variable if the independent variable (either standardized-continuous or binary) increases by a rational one unit. In addition, standardizing the continuous variable made the interaction effect more interpretable.

Results

Table 1 below show the results of regressing course grade (i.e., dependent variable) on OER, standardized previous GPA, standardized course difficulty, gender, Pell-eligibility, and the interaction between OER and standardized course difficulty (i.e., independent variables and the interaction term). The multiple R equals 0.446 with a coefficient of determination ($R^2$) of 0.199, which indicates 19.9% of the overall variance in the outcome, course grade, can be explained by the list of independent variables included in this study. The overall model is significant [$F (6,15,626) = 646.163, p < 0.0001$]. The zero-order correlation of OER with course grade was 0.025 which was significant ($p < 0.05$). However, in the presence of all the other predictors, OER was not a significant predictor of course grade ($B = 0.025, \beta = 0.005, p = 0.469$). All other predictors in the model were significant. Previous GPA is the strongest predictor ($B = 0.605, \beta = 0.410, p < 0.001$) and accounts for 16.6 percent of the variance in course grade [$semi-partial \ coefficient (0.408) squared = 0.166$]. The unstandardized coefficient of 0.605 means that there was a projected 0.605-point increase (in a 5-point grade scale) in student course grades with every unit (i.e., 1 SD) increase in student previous GPA, holding other predictors constant. Importantly, the covariate of standardized course difficulty was significant in the presence of the other variables ($B = -0.349, \beta = -0.169, p < 0.001$); that is a predicted decrease of 0.349 point in student course grades with every unit (i.e., 1 SD) increase in course difficulty while holding other predictors constant. This pattern is also consistent with the zero-order correlation between course difficulty and course grade ($r = -0.159$). Reasonably, the coefficient was negative, meaning that course grades tended to be lower as course difficulty increased. Standardized course difficulty was based on the aggregated failure rate of each course which was based on student course grades. However, because the standardized course difficulty was aggregated across multiple sections for each course and the student course grade was based on individual performance, the zero-order correlation between them was not problematic with only one percent shared variance ($r = -0.138, r^2 = 0.019$). This strategy to estimate course difficulty is recommended as there do not appear to be issues with multicollinearity but does require a large sample of sections and courses.
Most importantly, the interaction between OER and standardized course difficulty was significant ($B = 0.248, \beta = 0.039, p < 0.001$). The positive valence of the interaction term indicates that although the general trend (main effect) is for course grade to decrease with increased standardized course difficulty, the presence of OER blunts the impact of standardized course difficulty on course grades.

**Table 1. Regression of Course Grade on OER, GPA, Course Difficulty, Gender, Pell and Interaction between OER and Course Difficulty**

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>$\beta$</th>
<th>$t$</th>
<th>$p$</th>
<th>Zero-order</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>2.713</td>
<td></td>
<td>150.104</td>
<td>$p = 0.469$</td>
<td>0.025</td>
<td>0.926</td>
</tr>
<tr>
<td>OER course</td>
<td>0.025</td>
<td>0.005</td>
<td>0.724</td>
<td>$p = 0.469$</td>
<td>0.025</td>
<td>0.926</td>
</tr>
<tr>
<td>Previous GPA</td>
<td>0.605</td>
<td>0.410</td>
<td>57.009</td>
<td>$p &lt; 0.001$</td>
<td>0.412</td>
<td>0.992</td>
</tr>
<tr>
<td>Course Difficulty</td>
<td>-0.349</td>
<td>-0.169</td>
<td>-22.354</td>
<td>$p &lt; 0.001$</td>
<td>-0.159</td>
<td>0.894</td>
</tr>
<tr>
<td>Male</td>
<td>-0.054</td>
<td>-0.019</td>
<td>-2.555</td>
<td>$p = 0.011$</td>
<td>-0.065</td>
<td>0.965</td>
</tr>
<tr>
<td>Pell Eligibility</td>
<td>-0.151</td>
<td>-0.053</td>
<td>-7.280</td>
<td>$p &lt; 0.001$</td>
<td>-0.040</td>
<td>0.982</td>
</tr>
<tr>
<td>OER X DIFF</td>
<td>0.248</td>
<td>0.039</td>
<td>4.943</td>
<td>$p &lt; 0.001$</td>
<td>-0.012</td>
<td>0.842</td>
</tr>
</tbody>
</table>

Figure 1 below illustrates the significantly different slopes of the OER sections versus the non-OER sections using standardized course difficulty to predict course grade. The plot in Figure 1 is at zero-order for simple visualization purposes. However, it is very similar to, and does not distort the image based on the plot of the predicted values that accounts for all the covariates in the model. As seen in Figure 1, the negative slope of the OER course is less severe than the negative slope of the non-OER courses.

**Discussion**

The purpose of this study was to examine potential interactions between course difficulty and OER adoption on student grades. In addressing this purpose, we controlled for several potential confounds as recommended in Clinton and Kahn's (2019) meta-analysis while examining 15,633 course grades across 35 different college courses. Specifically, we controlled for self-reported gender, Pell eligibility, and importantly, previous academic performance. There was indeed an interaction...
between OER use and course difficulty on student grades in that OER adoption appeared to lessen the negative relationship between course difficulty and final grades.

To address this study’s purpose, we calculated the unique measure of course difficulty based on the proportion of students who earned a D or F in each course. Because the study included multiple sections of many courses over multiple semesters, the calculation of current failure rate is arguably logical and stable. While course difficulty was ultimately dependent on individuals’ course grades, the aggregation of failure rate across many sections and semesters did not result in undue multicollinearity, likely because of the large sample size involved.

The most important novel finding in this study, however, is the significant interaction between course difficulty and OER. The interaction term emerged in the presence of controlling for several potential confounds which typically “consume” available variance in multiple regression models predicting course grades. Finally, the interaction term emerged in the presence of a most powerful predictor, past student achievement. One potential explanation for this is that students’ need for course materials to perform well in a course may increase with course difficulty. In other words, it is possible that students are able to manage in less difficult courses without access to course materials, but for the more difficult courses they need support beyond what is pro-
vided by attending class and other freely available resources. This explanation is not something we are able to specifically test in our dataset but is supported by previous research findings in which students reported that not being able to afford course materials had negative academic consequences (Florida Virtual Campus, 2018). The access hypothesis applies here in that the students in difficult courses who may have needed course materials, but perhaps could not afford commercial materials, benefited from access to OER (Grimaldi et al., 2019). Moreover, the findings from this study indicate that one particular context—course difficulty—may potentially explain the variability in study finding’s in Clinton and Kahn’s (2019) meta-analysis.

Finding that OER blunts the expected negative main effect of course difficulty on course grades is very hopeful. Whatever conditions exist in courses (instructor rigor, workload, speed of instruction, concreteness or abstractness of content, match between student interest/aptitude and content, instructor experience and effectiveness, or any other predictors) were subsumed parsimoniously, empirically, and quantitatively in the aggregated course failure rate. No causal claims are made, but prediction is powerful enough to justify gambling that OER used in historically difficult (higher failure rate) courses might blunt the negative trend. Certainly, the trend was not reversed. Difficult courses still tend to result in generally lower grades, but the presence of OER might make that phenomenon less so with zero cost to students.

The difficult courses are by definition more challenging for students. In addition to OER use, other pedagogical interventions may be considered in future investigations in order to promote student learning in difficult courses, such as collaborative learning, providing more formative feedback to students, or promoting student motivations in the course.

While the zero-order correlation between OER and course grade was positive and significant (due to the large sample size), its beta-weight in the overall model was not significant. Controlling for gender, Pell eligibility, previous academic success and course difficulty diminished the weak positive association between OER and student outcomes. Even so, the zero-order result, as weak as it was, and the null result in the overall regression model still support the use of OER. This is not necessarily because of improved student achievement but on the grounds that student achievement using OER is on par with student achievement using traditional textbooks with zero costs to students. This null finding is the most frequently reported outcome (see Hilton 2016, 2019). OER produces similar results at diminished financial costs to our most financially vulnerable students.

Conclusion

Previous research findings have shown that OER provide students with similar learning outcomes as commercial materials at a greatly reduced cost (Clinton & Kahn,
However, the efficacy of OER based on allowing students access to materials likely varies by context such as course, institution, and student characteristics (Grimaldi et al., 2019). In this study, we examined the potential context of course difficulty and found an interaction with OER use on course grades. Grades declined less with course difficulty when OER were used compared to when OER were not used. These findings are useful for instructors and institutions who may be considering OER adoption or methods of improving student grades in difficult courses.

References


