

Assessing Outcomes Across On-line and In-Class Instruction

William T. Bianco

Department of Political Science
Indiana University

wbianco@indiana.edu

This proposal is to compare educational outcomes in traditional and on-line versions of a popular political science class, Introduction to American Government. This comparison addresses a fundamental question about on-line classes: do students learn as much (or more) from these classes as they do in classes that meet face-to-face?

The findings of this research have two important implications. First, they contribute to the wider debate over the value of on-line classes. While there has been much discussion at about various on-line instructional technologies, very little attention has focused on actual educational outcomes. Within the limits of this study (a particular course, instructor, and implementation of on-line), that is exactly what I propose to do.

A second goal is to improve the Political Science department's course offerings. Currently Intro is offered three times during the academic year, accounting for over a quarter of POLS enrollments. The plan is to begin advertising the summer on-line class more widely to boost enrollments, including would-be transfers students at other IU campuses, and, potentially, students throughout the US. Before making this step, it is critical to determine what students get from the on-line class – particularly given the course is a prerequisite for many classes in political science and is required for several other majors on campus.

The core of this project is a classic field experiment, where the same course content (including an e-text delivered via Courseload) and examinations are used for traditional and on-line classes. Half of the research has been completed: during Spring and Summer 2013, I taught the Intro class three times: once as a large sectioned class, once as a small lecture, and once on-line. The three classes had almost the same content, instructional materials, and exams. In effect, everything held constant except the method of instruction, allowing exam scores to be used as direct comparison of educational outcomes (results are presented later). I am currently repeating the same 3-class sequence in Spring – Summer 2014 to generate additional data for analysis.

The most important finding from initial analysis is that student performance in the on-line class lagged behind performance in semester and summer classes taught face-to-face.

Moreover, the effect is quite substantial compared to other factors affecting student performance.

Intro American Government: Three Implementations

As I teach it, the Intro class emphasizes mechanism – the central premise is that nothing in politics happens by accident, that outcomes are the result of the choices people make. The goal is to understand American politics by examining these choices and their consequences. The secondary emphasis is on facts, both to provide basic information about American politics and to demonstrate that there is evidence behind the theories and arguments presented in the class. The class is ruthlessly contemporary, emphasizing the political issues, campaigns, and controversies that are part of students' everyday life.

The standard (semester) version of the class enrolls about 350 students, and was taught in a two lecture, one section format. When taught in the 4-week summer session, the class meets for two and a half hours per day and enrolls about 50 students. The on-line class was also taught in a 4-week summer session, with daily work plans, on-line resources such as practice quizzes linked to the text and definitional flash cards, additional writing and exam prep assignments, and frequent chat sessions with the instructor and an associate instructor.

The text for the classes is *American Politics Today*, which I have coauthored and is in its 3rd edition, and is assigned for the classes as an e-text on Courseload. The text brings with it an extensive text bank, and on-line resources including videos.¹ These resources are the foundation of the implementation of the on-line class; they also facilitated the comparative assessment of student performance across the three classes. The on-line class covered slightly fewer chapters/topics – 15 rather than the full 17 in the semester and summer class.²

The test-bank was also used to generate a measure of baseline information: at the beginning of each chapter, students would take an ungraded pretest comprised of multiple-choice questions from the chapter's test bank. This pretest provided measure of the student's baseline knowledge – what they knew about the topic before being exposed to readings, lectures, and other instructional content.

Exams for the classes were a combination of 30 multiple-choice questions and short essays (students are given 3 questions and must write on two). The initial analysis focuses on the multiple choice questions, as I was unable to secure the same AIs for the summer classes, thus if there was some difference in scores, it would be difficult to determine whether the differences were real or an artifact of new graders.

The format of the exams varied. In the semester class, four exams were conducted in-class class, while one was an on-line open-book, timed exam taken through Oncourse. During the summer class, two exams were in-class and three were on-line. All of the on-line exams were taken on-line. Apart from these differences, the exams were almost identical – in

¹ A description of the book including sample chapters can be on the publisher's (W.W. Norton) web site at <http://www.wwnorton.com/showcase/AmericanPoliticsToday/>. The various online resources for the text are at <http://www.wwnorton.com/college/polisci/american-politics-today3/full/welcome.aspx>

² The structure of this class reflects the instructor's participation in a CITL Workshop in Summer 2011.

particular, for each portion of the class, exams for the semester, summer, and on-line class all had the same 30 multiple choice questions.³

Assessing Student Performance: A Field Experiment

The three-class sequence of Intro courses that I taught last year was designed to create a classis field experiment, to measure the impact of different methods of instruction by keeping course content and examination content constant and manipulating the variable of interest, the type of class. The following table summarizes the similarities and differences between the classes:

| | Spring 13 Lecture | Summer 13 Lecture | Summer 13 On-line |
|------------------------|--|-------------------------|--|
| Duration | 14 weeks | 4 weeks | 4 weeks |
| Weekly Format | 2 lecture, 1 section | Daily lecture | On-line |
| Exams | 4 in-class 1 on-line | 3 in-class 2 on-line | 5 on-line |
| Chapters (17 total) | 17 | 17 | 15 |
| Assignments | - | - | 8 Writing Assignments Review Quizzes On-line Flash Cards |
| Enrollment | 310 | 50 | 45 |
| Common Elements | Textbook, On-line Resources, Exam Content, Pre-Tests | | |

The “data” from this experiment consists of exam scores – the unit of analysis is a student’s score (out of 100) on one of the multiple-choice exam sections. In all, across the three classes, I collected over 2000 observations, each a single student’s performance on a given exam in a particular class. The preliminary analysis of this data was accomplished using a multiple regression, where the dependent variable is a student’s score on a particular exam.⁴ The following independent or explanatory variables were used:

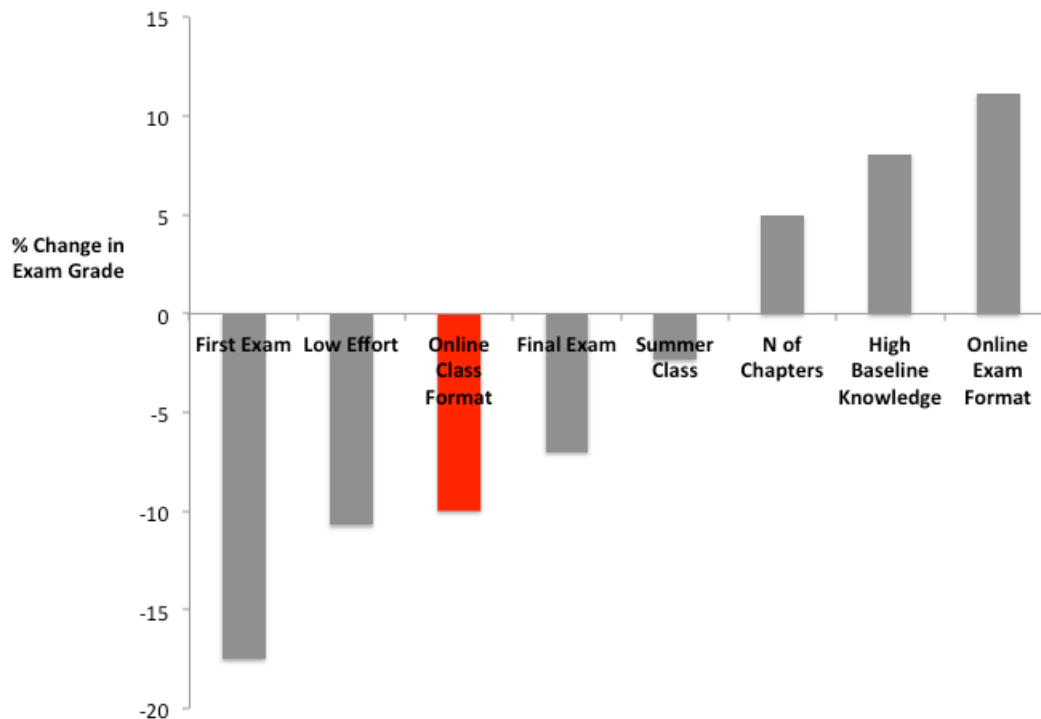
- Whether the exam was the first one in the class
- Whether the exam was the final exam in the class
- Whether the exam was conducted on-line or in-class

³ For 2 exams in the on-line class, the set of questions were slightly different because the section of the class covered 1 less chapter.

⁴ Full details of the analysis are omitted due to the space limitation of this proposal. I would be happy to provide any additional information on request.

- Whether the class was taught during the semester or in the summer
- Whether the class was taught on-line or face-to-face
- The number of chapters students were responsible for on the exam
- Student effort, measured using the percentage of completed pretests and assignments
- Baseline knowledge, measured using the pre-tests

Using the regression parameters, I calculated the predicted change in student exam performance as a function of these different explanatory variables. The figure below shows the direction and magnitude of these predicted changes, expressed as a percentage of the baseline exam score.



Some of these predicted changes are not surprising – students who enter the class knowing more do better; students who fail to turn in assignments do worse. Allowing an open-book exam increases scores by about 10%. Performance is slightly worse on final exams, presumably given the press of other work. Somewhat surprisingly, the largest negative effect is performance on the first exam, a point that will be addressed in the design of the 2014 classes.

The surprising finding of the analysis is that exam performance in on-line classes is significantly worse than in the other class types – about 10% worse. The question is, does this gap reflect something about the design of the class or is it a function of some extraneous factor that is currently not accounted for in the analysis? Are the same differences apparent in essay grades? Can a different instructional format or course design for the on-line class narrow the gap in achievement?

Moving Forward: Plan of Work

One factor that the current analysis does not consider is a student's class standing. In particular, it appears that the first implementation of the on-line class attracted many graduating seniors who need one additional class to fulfill credit hour or degree requirements.⁵ It is possible, then, that the observed difference between the classes may reflect the characteristics of the student body rather than the difference in method of instruction. To address this question, I will have research assistants add information on class standing into the analysis dataset, and include these variables in my analysis. We will add data on students' hometowns, to determine whether being a non-native English speaker is a significant influence on exam outcomes (anecdotal information suggests that it may be).

The second and third pieces of new analysis will build on my second round of teaching three versions of the intro class. In particular, I will structure my Spring 2014 class so that a smaller group of two AIs grades all of the essays for two of the exams. The same AIs will grade the same questions in the summer and on-line class using the same rubric, allowing a direct comparison of educational outcomes on this section of the test.

Budget

I request \$2000. All of the money (100 hours * 20/hour) will fund an RA who will hand-code information on students' class standing and hometown (about 800 cases) and combine the Excel grading spreadsheets into a STATA dataset that is organized with student-exam as the unit of analysis. This RA will share a coauthorship on the paper or papers that emerge from the data analysis. (I have already received IRB approval for this study.)

Prospects for External Funding

The research design developed for POLS 103 can easily be extended to other classes – for example, the department's POLS 205 (Statistical Methods) class. An on-line version of this class has been implemented, and received fairly poor reviews from students. What we do not know is how performance in 205 varies across the on-line and in-class versions of the class. The same question arises for other on-line classes across the College. We are at the initial stages of developing a cross-department analysis intended to test different on-line methods to determine best practices – or establish that, at least for now, that IU should focus on traditional face-to-face instruction.

⁵ The same might be true for some students in the other two classes, but preliminary work suggests the percentages are much smaller.