Supplemental Online Resources Improve Political Methods Education

Marco Alcocer*[1], Leonardo Falabella[1], Alexandra Lange[1], Nicholas Smith[1], and Maureen Feeley[1]

[1]University of California, San Diego

January 2023

Abstract

Mastery of knowledge and skills taught in introductory research methods/statistics courses is increasingly important for undergraduate political science majors, yet many students struggle in these courses. Asynchronous Online Supplemental Instruction (OSI) may offer a time- and cost-efficient means of supporting these students. However, we know little about the efficacy of these resources in general, or specifically in political methods education. This paper introduces an original OSI resource, "Foundations of Quantitative Research in Political Science," which includes content on key concepts and application of statistical methods to political problems. Utilizing a pre-registered within-subject experimental design, we find that access to this resource significantly improves student performance in an introductory political methods course at a large public university. To our knowledge, this study is the first to estimate the causal effects of OSI in political science, and specifically in a political methods course. Moreover, the research design itself is relatively easy to replicate, fair, controls for student-specific characteristics, and generates a large number of observations, which may be especially useful to those who teach small classes and are otherwise constrained by small N analyses. **Keywords:** Methods education; online supplemental instruction; experiment.

*Corresponding author. Email: m2alcoce@ucsd.edu.

Introduction

The American Political Science Association (APSA) reports that 82.4% of undergraduate political science and government programs in the United States offer a research methods/statistics course as part of their core curriculum (Davis, McGrath, and Super 2019).[1] Yet, research finds that many political science majors struggle in these courses (Buchler 2009; Adriaensen, Kerremans, and Slootmaeckers 2015; Slocum-Schaffer and Bohrer 2021; D. S. Brown, Bryant, and Philips 2022; Harden and Esarey 2022). This, in turn, may negatively impact student retention and time-to-degree in political science programs. Moreover, unless students master the knowledge and skills introduced in these foundational courses, they are disadvantaged in upper-division coursework and in co-curricular and postgraduate opportunities (e.g., research, internships, employment, and postgraduate study) that increasingly require these skills.

To support student learning in challenging courses, some universities have developed optional face-to-face supplemental instruction (SI) that are regularly scheduled to complement course instruction during the semester, and research has found these to be largely successful (Arendale 2020; Bowman, Preschel, and Martinez 2021; Dawson et al. 2014; Ogden et al. 2003; Spaniol-Mathews, Letourneau, and Rice 2016; Moradi et al. 2018). However, providing these resources for students is cost intensive as each semester they require, among other resources, SI leaders, classrooms, and active faculty or staff oversight.

To avoid these costs, asynchronous online supplemental instruction (OSI) may offer a time- and cost-efficient means of supporting students in challenging classes, given that it does not require additional instructional support beyond its initial construction. Yet, the literature has not convincingly addressed whether asynchronous OSI improves student learning in general, nor has it examined its effects in political science specifically. Existing studies to date on OSI remain limited by methodological challenges including small sample sizes (Moradi et al. 2018) and selection bias (Paloyo 2015; Bowman, Preschel, and Martinez

[1] APSA. 2019. "2017-18 APSA Departmental Survey: Enrollments and Curriculums." This survey was administered online from June 6 to October 15, 2018 to 1,263 departments at four-year colleges and universities offering degrees in Political Science and Government in the U.S., with 383 departments (30.3%) responding.

2021), with the only experimental study, to our knowledge, having a small sample size (Moradi et al. 2018). Furthermore, this research has focused almost exclusively on STEM courses, with no study to date having studied the impact of OSI in political science or political science methods courses.[2]

This study introduces an asynchronous online supplemental instruction (OSI) resource, "Foundations of Quantitative Research in Political Science," specifically designed to support student learning in introductory political science methods courses. The resource is comprised of ten interactive modules that include: (1) a brief textual overview of the module's content and learning outcomes; (2) one or more short videos (approximately seven to ten minutes) that embed concepts within a motivating political problem, focus on addressing misconceptions, and explain the logic underlying concepts; (3) a brief textual summary or "recap" of the module with links to additional supplementary materials for students to "dig deeper," (4) a "knowledge check"—interactive quiz(zes) with feedback to assess how well students understood the module,[3] and (5) a "reflection" opportunity that invites students to identify what they have learned and provide feedback on the module.

To evaluate the impact of the modules on student learning, we conducted a pre-registered double-blind within-subject experiment, randomly assigning all students in the course to a subset of the modules. We then designed midterm exam questions to evaluate students' mastery of specific concepts and applications addressed both in course materials and by a subset of OSI modules. That is, each student taking the exam answered questions for which they both did and did not have access to modules, so that no student was unfairly advantaged by the research design. In this way, each student served as their own control group, allowing us to control for individual student differences, while also ensuring that all students had access to an equal number of randomly selected modules.

We find that access to the OSI modules significantly improved student performance

[2]To our knowledge, only Warren and Tonsetic (1998) and Ogden et al. (2003) analyze traditional SI in political science courses, but neither analyze OSI, methods courses in political science, or provide causal evidence of their effects on student learning. Other studies, such as Yue et al. (2018) and Skoglund, Wall, and Kiene (2018) include political science courses in their sample, but do not report results for this subsample of courses.

[3]Some modules include multiple "knowledge checks"/quizzes, and students are encouraged to interact with the modules as often as necessary until mastery is achieved.

on exam questions, with larger effects for students who engaged more deeply with the resources. Specifically, we find that simply having access to the modules increased exam question grades by 3.8 percentage points. For students who viewed the modules, exam question scores increased by 5.3 percentage points, and if students answered a module quiz, scores increased by nearly 11 percentage points. We additionally tested for heterogeneous effects on students based on pre-treatment cumulative GPA and underrepresented status, and find no evidence of statistically significant differential effects. That is, the modules benefited all students equally, regardless of prior academic achievement (as measured by pre-treatment cumulative GPA)or minority status. Finally, we also collected information on students perceptions of the modules through a voluntary post-experiment survey. Although the sample size was small, and selection bias was likely present, an overwhelming majority of students reported that the modules helped them learn and achieve higher grades, and that they would recommend the modules to other students.

This study makes three main contributions. First, while the positive impacts of traditional SI on student learning have been well established, much less is known about the impacts of OSI, which have become increasingly common, especially since the global pivot to remote learning in the wake of the Covid-19 pandemic. This study contributes to this research by providing causal evidence that asynchronous OSI can positively, and significantly, impact student learning in challenging courses. Second, to our knowledge, this study is the first to utilize a within-subject experimental design to assess the impacts of OSI. The advantages of this design are that it is more fair to students than many alternatives; it addresses endogeneity and selection bias concerns in prior research; and it generates a large number of observations, thus increasing statistical power. It is also relatively easy to replicate. Thus, the study contributes to the broader scholarship of teaching and learning (SoTL) literature on equitable and effective research design to assess student learning in this area. Finally, the study also develops a novel asynchronous OSI resource specifically to support student learning in introductory quantitative methods courses in political science, with demonstrated positive and causal impacts on student learning. In so doing, the study also contributes to political science education research on teaching undergraduate quantitative

methods, an area of political science that has become increasingly important for student success in, and beyond, the major, yet an area where many students struggle.

Course Background

The Political Science Department at UC San Diego (UCSD) requires that all of its undergraduate majors take a ten-week introductory political methods course, POLI 30, Political Inquiry. This course introduces students to fundamental tools of political inquiry, including quantitative data, statistical software, probability theory, measurement, inference, research design, hypothesis testing, linear regression, and other basic statistical methods. POLI 30 is offered every quarter (Fall, Winter, Spring, Summer) and typically enrolls between 200 and 250 students each quarter during the academic year. The course is taught by multiple instructors, resulting in some variation in course content, but core concepts and learning outcomes remain consistent across instructors. In each offering of the course during the academic year, students meet twice a week for 50-minute lectures led by the course instructor, with an additional weekly meeting (50 minutes) in smaller discussion sections (approximately 30 students per section) led by graduate student TAs.

POLI 30 has double the average DFW (letter "D" grade, failure, withdrawal) rate of all courses in the political science curriculum. The literature has found that students struggle in political methods courses for a variety of reasons. First, students often arrive with varying levels of prior exposure to prerequisite mathematical concepts as well as "fixed mindsets" or preconceptions about their ability to succeed (Buchler 2009; Adriaensen, Kerremans, and Slootmaeckers 2015). Second, the type of learning required in quantitative courses is often significantly different from other courses in the curriculum in that learning tends to be both more linear and cumulative (Buchler 2009). Third, students often mistakenly assume they can succeed in these classes by memorizing statistical formulae, rather than focusing their efforts on understanding underlying principles and logic (Buchler 2009). Finally, as has been well established by constructivist theories of learning, mastery requires frequent low-stakes opportunities to practice new knowledge and skills with prompt feedback (P. Brown, H.

Roediger, and McDaniel 2014). Yet, many of these courses have high student-to-faculty and student-to-TA ratios, which constrains instructors' and TAs' ability to provide the frequent individualized feedback we know students need to succeed.

The authors have also experienced first-hand how many students struggle in the course, and at a large public university like UCSD, these challenges are often exacerbated by the fast-paced nature of the ten-week quarter system, as well by high student-to-faculty (typically between 200 and 250 students) and student-to-TA ratios (about 60 students per TA).

Foundations of Quantitative Research in Political Science

To support student learning in POLI 30, we created an asynchronous OSI resource, "Foundations of Quantitative Research in Political Science." This resource includes ten modules, with the first introducing students to the supplemental materials and the scientific method, and the remaining nine focusing on specific course topics that students have historically struggled with most.

As noted above, each module is introduced by a brief textual overview of the module and a clear statement of learning outcomes, which aligns with research on effective pedagogy (Ambrose et al. 2010; Nilson 2016). Following previous research on teaching political methods (Buchler 2009; Adriaensen, Kerremans, and Slootmaeckers 2015) and insights from constructivist theories of learning (Ambrose et al. 2010; National Research Council 2000), the modules were designed to introduce fundamental concepts in an intuitive way, rather than a more math-heavy or highly technical approach. The modules introduce real-world political problems to motivate the logic underlying key course concepts, address common misconceptions, and provide examples that highlight the political relevance and potential applications of concepts. Figure 1 displays the user interface of one of the modules and a screenshot from an instructional video.

Building on research from cognitive and neuroscience (National Research Council 2000) and multimedia learning (Mayer 2014), as well as guidance from educational and technology

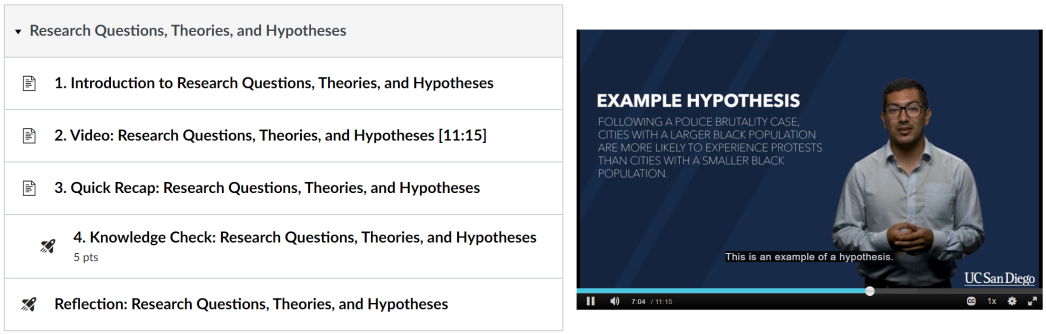

Figure 1: User interface of the OSI modules as seen on Canvas. Left screenshot is from the Canvas menu showing the Research Questions, Theories, and Hypotheses module. Right screenshot is from the Research Questions, Theories, and Hypotheses video.

specialists at UCSD,[4] videos were kept relatively short (approximately seven to ten minutes) and were filmed in UCSD's professional studio. Each video is accompanied by a "quick recap" that briefly summarizes the main learning objectives of the video as recommended by prior research in this area (Ambrose et al. 2010; P. Brown, H. Roediger, and McDaniel 2014; Moradi et al. 2018; National Research Council 2000; H. L. Roediger and Karpicke 2006).

Module summaries are then followed by "knowledge checks" (i.e., quizzes) that provide students with opportunities to check their mastery of the module's concepts. "Knowledge check" questions were designed to probe common misconceptions, and question banks were created to enable students to take multiple self-tests with immediate feedback on their level of mastery. The pedagogical value of providing students self-testing opportunities with prompt formative feedback is well established in the literature on human cognition (Ambrose et al. 2010; National Research Council 2000; P. Brown, H. Roediger, and McDaniel 2014; National Research Council 2000; H. L. Roediger and Karpicke 2006), but as discussed above, is often challenging to implement in large-enrollment courses such as POLI 30.

The final section of each module asks students to reflect on their learning and to provide feedback on the module itself—both of which are pedagogical practices that research has

[4]April Cha, Instructional Designer, UCSD; Caryn Neiswender, Educational Specialist, UCSD: Galen Davis, Senior Educational Technology Specialist, UCSD; Seth Marshburn, Senior Production Director, UCSD.

demonstrated deepen learning, improve retention, and build metacognitive skills (Ambrose et al. 2010; National Research Council 2000).[5] Beyond these common elements of each module, the modules themselves are generally ordered from more foundational to more complex concepts. Depending on their needs, students are free to engage with modules in whichever order they choose, however.

To ensure that students felt comfortable navigating the resources, we chose to host the modules on Canvas,[6] since Canvas is the default learning management system (LMS) used at UCSD. In typical offerings of POLI 30, students have access to all ten modules on a single Canvas course via a "join" link during the first week of the quarter. Once students join the OSI course, the modules populate their home Canvas page as a separate course.

Impact Evaluation

To evaluate the impact of the modules on student learning, we conducted a pre-registered double-blind experiment using a within-subject design. Specifically, we randomly assigned each student's access to a subset of the OSI modules. We then designed midterm exam questions to evaluate student knowledge, taking care to assess understanding of course content that was addressed *both* in lectures and in the online supplementary modules, so that no students were unfairly advantaged by the design.

Each student taking the exam answered questions for which they did and did not receive treatment. For any given student, an untreated question tested their knowledge on a topic that was presented in standard course material (textbook, lectures, discussion sections, and office hours), but for which they did not have access to the relevant OSI module. By contrast, a treated question tested a student's knowledge on a topic presented both in standard course

[5] Questions include: (1) "What are your main takeaways from this module?" (2) "What questions remain for you?" (3) "Do you have any feedback about this module to share? How can this module be improved?" and (4) "Please rate your agreement with the following statements," with statements including (a) "I found this module helpful," (b) "The video(s) in this module improved my understanding of this module's content," and (c) "The knowledge check(s) in this module improved my understanding of this module's content." Questions 3 and 4 were added after we conducted our experiment, but we include them here as a recommendation for obtaining low-effort student feedback, as students may not want to spend the time to answer open-ended questions.

[6] Canvas is an online learning management system (LMS) developed by https://www.instructure.com/ and used by universities, educators, and students to manage and access course content.

Table 1: Access to modules by treatment group. All groups had access to the Introduction and Research Questions, Theories, and Hypotheses modules. Each group was then randomly given access to two of the four remaining modules.

	Treatment Group					
Module	**1**	**2**	**3**	**4**	**5**	**6**
Introduction	x	x	x	x	x	x
Research Questions, Theories, and Hypotheses	x	x	x	x	x	x
Introduction to Variables	x	x	x			
Confounding and Intervening Variables	x			x	x	
Research Design		x		x		x
Introduction to Inference			x		x	x

material and the OSI modules randomly assigned to them.

Our sample consisted of undergraduate students at UCSD who were enrolled in the course during spring quarter 2021, excluding those who opted out or did not take the midterm exam. All students who consented to participating in the experiment were assigned to treatment, including students who enrolled late. For the purposes of the experiment, we created six different Canvas courses, one for each treatment group, and all enrolled students were randomly assigned to one of these six treatment groups.[7]

Two modules were available for every treatment group: the module that introduced the resources and a module on research questions, theories, and hypotheses. (These contain foundational knowledge and thus are crucial for understanding subsequent modules.) Each treatment group was also provided access to two additional modules randomly selected from a set of four (Introduction to Variables, Confounding and Intervening Variables, Research Design, and Introduction to Inference), with each treatment group having access to a different set of modules. Table 1 provides a complete description of access to modules by treatment group.

Students were given access to their respective Canvas pages during the first week of the course and were notified about materials both via email and in class. Students were also aware that the effectiveness of the resource was being tested, but not aware of design details, including their treatment group. Students were sent invitations to their Canvas page multiple times in the weeks leading up to the midterm exam. It is important to note

[7]Students were introduced to the study and the opt-out IRB form during week 1 of the quarter.

that students were encouraged, but not required, to use the supplementary materials, and no additional incentives (e.g., extra credit points) were awarded for their use. Moreover, because we controlled access to the Canvas pages, students could not access the Canvas pages for other treatment groups.

A particular strength of our design is that it was double-blind. Neither the students nor the instructors (professor and TAs) were aware of the treatment assignment of each student. We also did not provide instructors access to the Canvas pages or OSI modules to prevent them from adjusting their teaching given the contents of the OSI modules.

Five weeks into the ten-week course, students took a midterm exam worth 20% of the course grade. The midterm exam (included in the SI Appendix) consisted of 18 questions, 13 of which are the focus of our impact evaluation. Four questions (1a-1d) asked about course content that was not covered by the supplemental resources, and one (2a) asked about content taught in the "Research Questions, Theories, and Hypotheses" OSI module, which was available to all six treatment groups. Consequently, these questions were dropped from our study. By limiting ourselves to this subset of 13 questions, we limit our analysis to questions that focused on content covered in the OSI materials and that were available to some treatment groups, but not others. Table 2 provides a complete description of how modules and questions overlap.[8]

The unit of observation in this study is student-question. Treated units include those questions answered by students who had access to a relevant OSI module. Untreated units include student answers to questions covered by OSI material but *not* available to them. The main advantage of this research design is that errors associated with differences in skill and effort between students are reduced, since each student acts as their own control group.

In total, there were 216 students enrolled in the course during the first week of the quarter. Of those, 189 students took the midterm exam. Forty-six students were dropped from the study either because they opted out or were minors at time of consent. Thus, our study includes 143 students and our student-question data set contains 1,859 observations (143 students $\times$ 13 questions). Treatment groups ranged in size from 18 to 28. We chose

[8]Question 3a asks about information covered by two modules, which should bias our estimates towards zero.

Table 2: Mapping between material covered in each module and material covered in each exam question.

	Q2			**Q3**		**Q4**				**Q5**			
Module	**b**	**c**	**d**	**a**	**b**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**
Introduction													
Research Questions, Theories, and Hypotheses													
Introduction to Variables	x	x	x										
Confounding and Intervening Variables				x	x								
Research Design				x						x	x	x	x
Introduction to Inference						x	x	x	x				

to evaluate the impact of the OSI modules on midterm exam performance so that students could have full access to the modules for the remainder of the quarter after they took the exam. The exams were graded by five political science PhD students using the Gradescope application, with each PhD student grading the same set of questions across all students.[9]

Several strengths of the research design are worth highlighting. First, it provides a more ethical way of testing the effects of OSI than alternative experimental designs since it ensures that all students receive some randomly determined subset of modules. That is, it was not the case that some students were given access while others were not. Second, the design allows us to measure the effects on specific learning outcomes, rather than effects on overall course or exam grades, enabling us to provide direct evidence that the modules affect the specific outcomes for which they were intended. Third, the within-subject design increases the number of observations since each student is observed multiple times, increasing statistical power. It also allows for the inclusion of student fixed effects, which absorb student-specific variance. If student performance *across outcomes* looks very different from one student to another, then this fixed effect becomes a good predictor of the outcome and leads to more precisely estimated estimates. Finally, the within-subject design allows us to better detect student-level heterogeneity across subgroups of students since the individual treatment effects can be estimated for each student.

[9]To ensure consistency, we created a detailed rubric for each question and trained each grader on the grading rubric. Questions 3a and 3b were graded by one of the co-authors of this study, and all remaining questions were graded by PhD students serving as TAs for POLI 30, none of whom were directly involved in the study. TAs were not aware of treatment assignments when grading. TAs also had no access to the online modules, which eliminates the concern that they would modify their teaching to fit the material covered by the modules. As a robustness check, we conducted our impact evaluation with a subsample that excludes questions 3a and 3b to account for potential sources of bias. All results are consistent with the main results.

OSI Availability

Our experiment was implemented as an encouragement design where the assignment to treatment was randomized but treatment compliance was not obligatory. As a result, non-compliance (students having access to but not using the OSI modules) is present, meaning we cannot estimate the average treatment effect (ATE). We therefore first estimate the intent-to-treat effect (ITT); that is, the effect of having access to the instructional materials on exam question scores. We use OLS linear regression to estimate the ITT.[10]

We estimate the regression using question-student level data. The regression model is:

$$Y_{iq} = \beta TreatmentAssignment_{iq} + \gamma_i + \lambda_q + \epsilon_{iq} \tag{1}$$

where q denotes each question and i denotes each student. β is the causal coefficient of interest. Y_{iq} denotes student performance on each exam question, which we measure using both percentages (0%-100%) and standardized scores. The treatment is a dummy variable indicating whether student i had access to supplemental modules addressing question q. γ_i are student fixed effects, and λ_q are question fixed effects.[11] The exam question fixed effects should absorb any differences in grading across exam questions as well as factors that affect each question equally across students.

Student fixed effects control for any time-invariant observable or unobservable student-specific characteristic. The within-subject design means that each student acts as their own control group, allowing us to estimate within-subject variation in outcomes when they received treatment versus when they did not. Finally, because our treatment is assigned at the student level, we cluster the standard errors at the student level for all models.

One concern may be spillover effects—students accessing OSI modules they were not assigned—which would violate identification assumptions and bias the estimates. We believe this is unlikely because students could not access OSI modules they were not assigned. Still, students may have shared their modules with other students. This also seems im-

[10] Non-compliance means that we cannot estimate the average treatment effect (ATE) of the modules.

[11] We do not include question fixed effects in models that measure the dependent variable as standardized scores.

probable because students would have needed either to share their confidential university username and passwords, or physically show the modules on their computers to other students. Because the experiment was conducted during the Covid-19 pandemic when courses were virtual, university facilities were closed, and many students were not living on campus, this seems unlikely. Even if spillover was present, this would mean students may have improved their scores for control questions, which would bias our estimates towards zero. In other words, even if spillover effects were present, they would bias estimates against finding treatment effects.

OSI Compliance

The ITT does not tell us the effect of students using the OSI resource on student learning, which is of great interest. By hosting the OSI modules on Canvas, we were able to collect information on treatment compliance—students engaging with the modules they were assigned. Yet, simply regressing module use on student outcomes would result in biased results because compliance is likely endogenous. Indeed, we find suggestive evidence of this: there is a positive correlation between pre-treatment GPA and engagement with the resource and suggestive evidence that underrepresented students engaged more with the resource, though neither of these differences are statistically significant.

Nevertheless, information on compliance does enable us to estimate the causal effect of the OSI resource on compliers, or the local average treatment effect (LATE) for compliers, also called the complier average causal effect (CACE), which estimates the average treatment effect (ATE) for compliers (Gerber and Green 2012, Ch. 5). In this study, the LATE is especially useful because it estimates the effect of the OSI resource on students who engaged with the resource.

We measure compliance in two ways (see Table 3 for descriptive statistics). First, we measure compliance by whether students viewed at least one OSI module at least once. This is a conservative measure that simply identifies whether a student was exposed to the treatment by looking at the modules available to them. Of the 182 students in the dataset, 122 students viewed at least one page. Second, we measure compliance by whether

a student completed at least one quiz. Of the 182 students, 93 completed at least one module quiz. This measure aligns more with our conceptual understanding of the treatment: actual engagement with the OSI modules. However, we interpret this second measure with caution as it categorizes students that used the OSI modules without answering quiz questions as noncompliers.

Since compliance is endogenous to student-specific characteristics, we estimate the LATE with an instrumental variables (IV) design (Gerber and Green 2012, pp. 157–160) where we regress the outcome on the treatment compliance, using treatment assignment as an instrument for treatment compliance. We use the two-stage least squares (2SLS) to estimate the LATE. The first stage is:

$$Treated_{iq} = \beta_1 TreatmentAssignment_{iq} + \gamma_i + \lambda_q + \epsilon_{iq} \tag{2}$$

where $TreatmentAssignment_{iq}$ denotes whether student i had access to online supplemental modules addressing question q. $Treated_{iq}$ denotes whether student i used online supplemental modules addressing question q. Results for the first stage show that the treatment assignment is a valid and strong instrument for whether students used the modules.[12]

In the second stage, we regress the outcome on the fitted values from the first stage:

$$Y_{iq} = \beta_2 \hat{Treated}_{iq} + \gamma_i + \lambda_q + e_{iq} \tag{3}$$

The coefficient of interest is Y_{iq}, which estimates the LATE. The reduced form for this instrumental variable is the ITT analysis from the previous section.

Heterogeneous Effects

To explore heterogeneous effects, also known as treatment-by-covariate effects (Gerber and Green 2012, Ch. 9) or conditional average treatment effects (CATE), we obtained de-identified student information on cumulative GPA prior to taking POLI 30 and minority

[12] We report the results of the first stage in the Appendix.

student status.[13] Table 3 shows descriptive statistics for these variables.

We estimate heterogeneous effects using linear regression:

$$Y_{iq} = \delta TreatmentAssignment_{iq} \times StudentChar_i + \gamma_i + \lambda_q + \epsilon_{iq} \tag{4}$$

where $StudentChar_i$ denotes the student-specific characteristic for which we have data (underrepresented or pre-treatment cumulative GPA). The interaction between this variable and the treatment assignment allows us to estimate the heterogeneous effects of interest, making δ the main coefficient of interest.

Results

We find that access to the OSI modules significantly increased student performance. Figure 2 visualizes the main results using exam question scores in percent as the outcome variable. On average, having access to the modules increased question scores by 3.8 percentage points, which translates to an increase of 0.14 standard deviations. While these estimates may appear small, they translate to about one third of a letter grade increase in the overall exam grade. That is, the average effects suggest that if students had access to the treatment for all four modules, their exam grades would have increased by 3 points out of 100. Further, point estimates from the LATE analysis of compliers suggest that questions scores were 5.3 points higher when students viewed at least one page related to that question, and nearly 11 percentage points higher when students completed at least one quiz. Finally, we do not find evidence that treatment effects varied across students based on their pre-treatment cumulative GPA or underrepresented status.

Table 4 shows the full regression output of the main results for both ITT and LATE estimations using both measures of the outcome variable. Models 1 and 4 show the results for OSI availability (ITT), while Models 2, 3, 5, and 6 show the second stage results for OSI compliance (LATE) using both measures of compliance (viewing a page and completing

[13] While we pre-registered the model specification presented here, we did not pre-register the particular student-specific covariates to be analyzed as there was uncertainty about which student-specific information would be made available to us by UCSD. The results should therefore be considered exploratory.

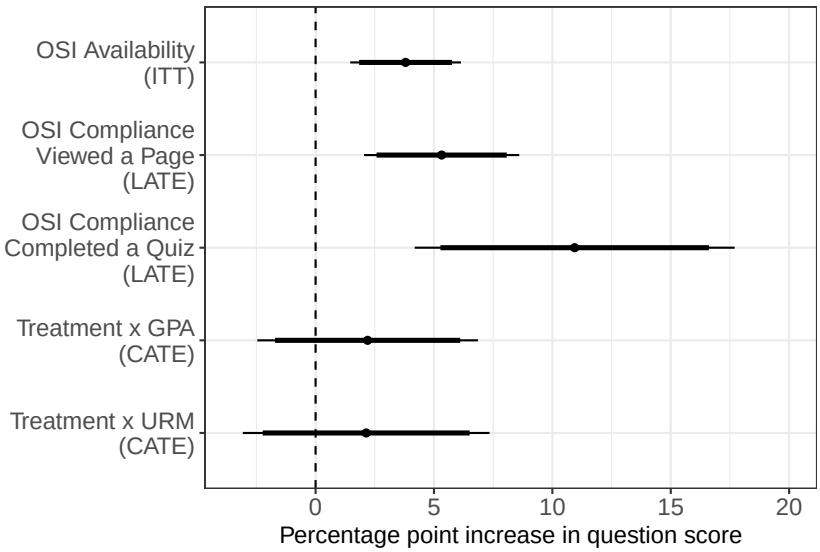

Figure 2: Plot shows point estimates from the main results using exam question scores in percent as the outcome measure with 90% and 95% confidence intervals represented with thick and thin lines, respectively. From top to bottom, plot shows effect of OSI availability (ITT) and OSI use (LATE) measured as viewing a page and completing a quiz, respectively, and effects of OSI availability conditional on student GPA and underrepresented status (URM) on exam question scores. Analysis includes student and question fixed effects. Robust standard errors are clustered at the student level. ITT = Intent to treat. LATE = Local average treatment effect. URM = student underrepresented status.

Table 3: Student-level summary statistics: modules use, GPA, exam question scores, and underepresented student status.

Variable	N	Mean	Median	Std. Dev.	Min	Max
Page views	143	29.25	24.00	27.93	0.00	141.00
Quiz completions	143	2.64	3.00	2.90	0.00	8.00
Incoming GPA	141	3.48	3.61	0.50	1.10	4.00
Exam Question Score	143	78.10	80.77	13.26	29.21	98.97

Variable	N	Perc.
Viewed $\geq$ one page	143	
... Yes	117	81.8%
... No	26	18.2%
Completed $\geq$ one quiz	143	
... Yes	73	51%
... No	70	49%
URM Status	131	
... Not URM	92	70.2%
... URM	39	29.8%

a quiz). For the IV-2SLS models, we estimate the first-stage cluster-robust F statistics and find that they are above conventional levels for a strong instrument, meaning that the treatment assignment is a valid instrument for both measurements of compliance ($F = 425.8$ for page views and $F = 116.3$ for quiz completion). The instrument being randomly assigned as part of an experimental design further validates its use.[14]

Table 4: Main results: Effects of OSI on student grades in exam questions. Models 1-3 use question grades in percent as outcome measure. Models 4-6 use standardized question grade as outcome measure. Models 1 and 4 estimate the ITT, models 2 and 4 estimate the LATE using module view as compliance, models 3 and 6 estimate the LATE using quiz taking as compliance.

	Question Score					
	Percent			Standardized		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment (OSI available)	3.801***			0.143***		
	(1.371)			(0.050)		
Compliance (viewed a page)		5.325***			0.198***	
		(1.870)			(0.069)	
Compliance (completed a quiz)			10.943***			0.418***
			(3.886)			(0.146)
Model	OLS	IV-2SLS	IV-2SLS	OLS	IV-2SLS	IV-2SLS
Student FE	Yes	Yes	Yes	Yes	Yes	Yes
Question FE	Yes	Yes	Yes	No	No	No
Observations	1,859	1,859	1,859	1,859	1,859	1,859
R^2	0.466	0.465	0.461	0.257	0.255	0.250

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors clustered by student in all columns.

As expected, the LATE estimates are larger than the ITT estimates,[15] and LATE estimates for compliance measured as quiz-taking are larger than those of compliance measured as exposure to the treatment. Figure 2 and Table 4 show that giving students access to OSI resources improved question scores (Models 1 and 4), and the effect was larger for students who viewed the resource (Models 2 and 5), and even larger for students who engaged with it (Models 3 and 6). This suggests that the OSI modules help students more the more they

[14] We report results for the first stage in the Appendix.
[15] The ITT estimates are the reduced form of the 2SLS-IV model.

engage with them.

Table 5 shows full regression output for heterogenous effects. We find that the direction of both interaction estimates are positive, perhaps suggesting that students with higher GPAs and underrepresented students benefited more from the modules. However, the estimates are not statistically significant at any conventional level. We therefore find no evidence that the modules had differential impacts on students depending on their GPAs or underrepresented status.

Survey Results

In addition to encouraging empirical results, we received positive feedback from students. We conducted a survey at the end of the POLI 30 course to elicit student views on the course and the supplemental modules. While the sample is relatively small, it suggests that the modules were well-received.

Of those who answered the survey, 17 out of 19 respondents reported using the resources. All 17 of these students reported that the resources were helpful (Yes/No question, "Did you find these resources helpful?") Further, on a scale of 0-5, where 0 indicates "strongly disagree" and 5 indicates "strongly agree," students generally agreed that the resources helped them get a better grade (mean = 4, range = 3-5); that they would recommend these resources to future POLI 30 students (mean = 4.53; range = 3-5), and that they would like to continue having access to the materials (mean = 4.63; range = 3-5).

Discussion and Conclusions

Research finds that many political science undergraduate majors struggle in introductory quantitative methods and statistics courses, yet these courses have become increasingly important not only for students' academic success, but also for 21st professional opportunities beyond the classroom – post-graduate training, internships, and employment. Prior research has established the positive impacts of traditional supplemental instruction (SI) for improving student learning in challenging courses, but less is known about the impacts of

Table 5: Heterogeneous treatment effects: treatment-by-covariate effects of OSI access on exam question grades given students' incoming cumulative GPA and underrepresented minority status. Models 1 and 2 use question grade in percent as outcome. Models 3 and 4 use standardized question grade as outcome.

	Question Score			
	Percent		Standardized	
	(1)	(2)	(3)	(4)
Treatment	−3.861	3.343**	−0.177	0.134**
	(9.543)	(1.570)	(0.353)	(0.059)
Treatment × GPA	2.197		0.092	
	(2.620)		(0.097)	
Treatment × URM		2.136		0.059
		(3.464)		(0.130)
Student FE	Yes	Yes	Yes	Yes
Question FE	Yes	Yes	No	No
Observations	1,833	1,703	1,833	1,703
R^2	0.469	0.465	0.258	0.255
Note:				*p<0.1; **p<0.05; ***p<0.01
	Standard errors clustered by student in all columns.			

online supplemental instruction (OSI). This paper presents strong evidence that OSI may be a cost-effective alternative to traditional SI to support student success in challenging courses, such as political methods. Specifically, the study finds that OSI resources significantly improved student learning in a large-enrollment introductory political methods course during spring quarter 2021, and that the benefits for student learning were greater the more deeply students engaged with the resources. Moreover, the study also finds that all students benefited equally, regardless of prior academic performance or minority student status.

Beyond contributing to the general literature on the learning impacts of OSI, and specifically to research on political methods education, to our knowledge, this study is the first to utilize a double-blind within-subject experimental design to assess the impacts of OSI on student learning. An important advantage of this research design is that does not unfairly advantage any student. In our experiment, all students were given access to an equal number of randomly assigned OSI modules, the learning outcomes of which were assessed on midterm exam questions designed to assess knowledge and skills addressed both in course

materials and select modules. Additionally, since each student serves as their own control group, the design addresses endogeneity and selection bias concerns in prior research on OSI. Finally, a within-subject design is relatively easy to replicate, and generates a large number of observations, thus increasing the statistical power of experiments. For these reasons, this design may be especially beneficial to scholars wanting to conduct scholarship of teaching and learning (SoTL) research in small classes. In this regard, the study contributes to the broader SoTL literature on equitable and effective research design to assess student learning.

Finally, the study contributes to political science education by developing a novel asynchronous OSI resource, "Foundations of Quantitative Research in Political Science," with demonstrated positive, causal, and significant impacts on student learning. Given evidence of the extent to which our majors struggle in political methods and statistics courses, and the growing importance of these courses for their academic and professional success in an increasingly data-driven world, our hope is that the OSI resources introduced here can be widely shared, replicated, or used in ways to inspire new projects to support student learning in this critical area of our undergraduate curriculum.

Acknowledgements

We thank the faculty who teach POLI 30—Dan Butler, Scott Desposato, Peter Galderisi, and Thad Kousser—for their support on this project. We especially thank Dan Butler for allowing us to conduct the experiment in his spring POLI 30 class, and Scott Desposato for allowing us to pilot the resources in his winter 2021 class. We are also indebted to Jasper Cooper for his insights and support on research design, and thank Claire Adida, Jasper Cooper, Scott Desposato, Peter Galderisi, Umberto Mignozetti, and Molly Roberts for their insightful and generous feedback on earlier drafts of the article. We are also grateful to the specialists at the UCSD Teaching and Learning Commons and Education Technology Services (April Cha, Galen Davis, Seth Marshburn, Robin Martin, and Caryn Neiswender) for making the OSI videos possible. Finally, we thank our undergraduate research apprentices (Anna Azaryan, Samuel Do, Cole Hernandez, Genessa Marquez, and Klayna Martínez-Fragosa) for their help designing and testing the OSI modules.

Competing Interests

The author(s) declare none.

Funding

Funding for this work was provided by the University of California San Diego's Course Development and Instructional Improvement Program (CDIIP) grant, AY 2020-21.

References

Adriaensen, Johan, Bart Kerremans, and Koen Slootmaeckers (Jan. 2015). "Editors' Introduction to the Thematic Issue: Mad about Methods? Teaching Research Methods in Political Science". In: *Journal of Political Science Education* 11.1, pp. 1–10. ISSN: 1551-2169. DOI: 10.1080/15512169.2014.985017.

Ambrose, Susan A. et al. (2010). *How learning works seven research-based principles for smart teaching*. San Francisco, CA: Jossey-Bass.

Arendale, David R. (Dec. 2020). *Postsecondary Peer Cooperative Learning Programs: Annotated Bibliography 2020*. Report. DOI: 10.13140/RG.2.2.26016.30729.

Bowman, Nicholas A, Stephanie Preschel, and Danielle Martinez (2021). "Does Supplemental Instruction Improve Grades and Retention? A Propensity Score Analysis Approach". In: *The Journal of Experimental Education*, pp. 1–25.

Brown, David S., Katherine V. Bryant, and Andrew Q. Philips (2022). "Integrating the Use of Statistical Software into Undergraduate Political Methodology Courses". In: *PS: Political Science & Politics* 55.1, pp. 210–215.

Brown, Peter, Henry Roediger, and Mark McDaniel (2014). *Make It Stick: The Science of Successful Learning*. Cambridge, MA: Harvard University Press.

Buchler, Justin (2009). "Teaching Quantitative Methodology to the Math Averse". In: *PS: Political Science & Politics* 42.3, pp. 527–530.

Davis, Megan, Erin McGrath, and Betsy Super (2019). "2017–2018 APSA Departmental Survey: Enrollments and Curriculums". In: *American Political Science Association*.

Dawson, Phillip et al. (Dec. 2014). "On the Effectiveness of Supplemental Instruction: A Systematic Review of Supplemental Instruction and Peer-Assisted Study Sessions Literature Between 2001 and 2010". In: *Review of Educational Research* 84.4, pp. 609–639. ISSN: 0034-6543. DOI: 10.3102/0034654314540007.

Gerber, Alan S. and Donald P. Green (May 2012). *Field Experiments: Design, Analysis, and Interpretation*. Illustrated edition. New York: W. W. Norton & Company. ISBN: 978-0-393-97995-4.

Harden, Jeffrey J. and Justin E. Esarey (2022). "Introduction to the Symposium on Teaching Political Methodology". In: *PS: Political Science & Politics* 55.1, pp. 203–205.

Mayer, Richard E. (Feb. 2014). "Incorporating Motivation into Multimedia Learning". In: *Learning and Instruction* 29, pp. 171–173. ISSN: 0959-4752. DOI: 10.1016/j.learninstruc.2013.04.003.

Moradi, Moein et al. (2018). "Enhancing Teaching-Learning Effectiveness by Creating Online Interactive Instructional Modules for Fundamental Concepts of Physics and Mathematics". In: *Education sciences* 8.3, p. 109.

National Research Council (2000). *How People Learn: Brain, Mind, Experience, and School: Expanded Edition*. Washington, DC: The National Academies Press. ISBN: 978-0-309-07036-2. DOI: 10.17226/9853.

Nilson, Linda B (2016). *Teaching at its best: A research-based resource for college instructors*. John Wiley & Sons.

Ogden, Peggy et al. (2003). "Supplemental Instruction: Short-and Long-Term Impact". In: *Journal of Developmental Education* 26.3, p. 2.

Paloyo, Alfredo R (2015). "A note on evaluating supplemental instruction". In: *Journal of Peer Learning* 8.1, pp. 1–4.

Roediger, Henry L. and Jeffrey D. Karpicke (Mar. 2006). "Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention". In: *Psychological Science* 17.3, pp. 249–255. ISSN: 0956-7976. DOI: 10.1111/j.1467-9280.2006.01693.x.

Skoglund, Kirk, Timothy J. Wall, and David Kiene (2018). "Impact of Supplemental Instruction Participation on College Freshman Retention". In: *Learning Assistance Review* 23.1, pp. 115–135. ISSN: 1087-0059.

Slocum-Schaffer, Stephanie A. and Robert E. Bohrer (2021). "Information Literacy for Everyone: Using Practical Strategies to Overcome 'Fear and Loathing'in the Undergraduate Research Methods Course". In: *Journal of Political Science Education* 17.sup1, pp. 363–379.

Spaniol-Mathews, Patricia, Lawrence E Letourneau, and Ethan Rice (2016). "The Impact of Online Supplemental Instruction on Academic Performance and Persistence in Undergraduate STEM Courses." In: *Grantee Submission* 2.1, pp. 19–32.

Warren, Bonnie Z. and Robert Tonsetic (1998). "Supporting Large Classes with Supplemental Instruction (SI)". In: *Journal of Staff, Program & Organization Development* 15.2, pp. 47–54. ISSN: 0736-7627.

Yue, Hongtao et al. (2018). "Supplemental Instruction: Helping Disadvantaged Students Reduce Performance Gap". In: *Journal of Developmental Education* 41.2, pp. 18–25. ISSN: 0894-3907.