Supplemental Materials for "A Student-Centered Approach to Undergraduate Research Participation"

Avital Livny University of Illinois at Urbana-Champaign April 9, 2023

Online Appendixes

A1 Sample of Multi-Semester Research Curriculum

A1.1 Semester 1 Coursework

Unit 1: Introduction to the "What" and "Why" of Research and the "How" of the Apprenticeship

- Assigned Readings: Course syllabus and project description; Codebook(s) for apprentice-ship task(s).
- Assignment (Coursework): In a short paragraph (and in your own words): what do you see as the key contribution of the research project? What aspect(s) are you most interested in yourself?
- Assignment (Apprenticeship): Review the codebook(s) for your assigned task(s). Watch the provided training video(s). If prepared to, begin work on the first training exercise.
- *Goal*: Offer students an introduction to the "what" and "why" of research as it relates to the PI's project, where the students will gain their hands-on experience. In addition, training in the task(s) students will be asked to complete as part of their apprenticeship. Training exercises are distributed, and junior apprentices are encouraged to ask lots of clarifying questions of each other and senior apprentices, as well as of the PI.

Unit 2: Identifying Patterns and Puzzles

- Assigned Readings: Maryann Barakso, Daniel M. Sabet and Brian F. Schaffner. 2014. Understanding Political Science Research Methods: The Challenge of Inference. New York: Routledge. "The Challenge of Inference." pp. 11-27.
- *Additional Resources*: At least two scatterplots (possibly with fitted lines) or other figures that represent a bivariate relationship related to the PI's project.
- Assignment: Evaluate the figures provided and see if you can identify a puzzle worth addressing. Write up a short description of what you see and what you find puzzling.
- *Goal*: Develop an understanding that different eyes are drawn to different questions, that there is no "right" or "best" one. This lesson can be linked back to the previous discussion of everyone's interest in the collaborative project(s) that are part of students' apprenticeships.

Unit 3: Engaging with Social Scientific Literature

- Assigned Readings: Janet Buttolph Johnson, H. T. Reynolds and Jason D. Mycoff. 2016. Political Science Research Methods. Los Angeles: CQ Press. "Introduction" (select one topic).
- Assignment: Choose one of the topics introduced by Johnson et al. Identify at least two research questions addressed in the existing social scientific literature. How were data used to answer each question? What do you find interesting about these questions and the answers uncovered by the data? What additional question(s) do they raise?
- *Goal*: Help reinforce the idea that research involves collaboration among scholars with diverse perspectives, that research ideas are rarely wholly new, and that, as researchers, we can (and should) build off others' contributions. Where disagreements occur, these should

be respectful and impersonal. Encourage them to find a way of respectfully engaging with existing work on a topic and identify a small contribution that they can make to the ongoing discussion.

Unit 4: What are Data?

- Assigned Readings: Paul M. Kellstedt and Guy D. Whitten. 2013. The Fundamentals of Political Science Research. Cambridge: Cambridge University Press. "Thinking About the World in Terms of Variables and Causal Explanations." pp. 7-15; Maryann Barakso, Daniel M. Sabet and Brian F. Schaffner. 2014. Understanding Political Science Research Methods: The Challenge of Inference. New York: Routledge. "Different Types of Data." pp. 86-89.
- Additional Resources: Screenshots of at least two (if not three) datasets, ideally related to the PI's project. Data should be at different levels of analysis, possibly the same data at different levels of aggregation (e.g., individuals, countries).
- Assignment: Evaluate the spreadsheets provided. For each, identify the unit of analysis and then pick out at least one pattern (across one or more dimensions) that you would want to explore. For the latter, explain your choice in a sentence or two.
- *Goal*: Clarify key terms like data point, data set, variable, observation, unit of analysis, etc. Make clear that different questions require different types of data. Also serves as a nice introduction to the types of coding decisions that researchers make in the data-generation process and how data can be transformed through aggregation and recoding.

Unit 5: The Data-Generation Process

- Assigned Readings: Robert Adcock and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." American Political Science Review 95 (3): pp. 530-32; Janet Buttolph Johnson, H. T. Reynolds and Jason D. Mycoff. 2016. Political Science Research Methods. Los Angeles: CQ Press. "The Building Blocks of Social Scientific Research: Measurement." pp. 121-25, 130-32, 143-50.
- Assignment: Take a background concept tolerance and offer a conceptual definition ("systematized concept"), a way of operationalizing it (into one or more "indicators"), and a possible measurement strategy (creating "scores for cases"). In a couple of sentences, consider the potential for measurement bias in your proposed strategy.
- *Goal*: Understand the important decisions that go into defining (i.e., conceptualizing) and observing (i.e., measuring) phenomena that are not observable (i.e., latent). Reinforce that the same phenomenon can be defined in different ways, and that many of these are reasonable, but all require an explicit justification. Similarly, highlight how the different ways of observing indicators of the same variable. Offer a brief introduction of measurement error both random and systematic and how it can influence the patterns that we (think we) see.

Unit 6: Engaging with Ready-Made Data

- Assigned Readings: Gerardo L. Munck and Jay Verkuilen. 2002. 1"Conceptualizing and Measuring Democracy: Evaluating Alternative Indices." *Comparative Political Studies*. 35 (1): pp. 5-34. (Read strategically!)
- *Additional Resources*: List of datasets commonly used by scholars in different subfields of political science.
- **Assignment:** Spend some time looking for an existing dataset that is about a subject of particular interest to you. Prepare some notes on the dataset the variables included, how the data were collected, what questions you think could be explored using it. In our class meeting, you'll be asked to introduce the dataset to the group.
- *Goal*: Encourage students to make explicit the choices that researchers made in conceptualizing and measuring the variables in the dataset, considering possible sources of systematic measurement error. Within the dataset, students brainstorm potential relationships between variables, beginning to think about their first hypothesis.

Unit 7: Assessing Bivariate Relationships

- Assigned Readings: Paul M. Kellstedt and Guy D. Whitten. 2013. *The Fundamentals of Political Science Research*. Cambridge University Press. "Bivariate Hypothesis Testing." pp. 145-67.
- *Additional Resources*: List of links to websites explaining the basics of generating cross-tabulations, calculating correlation coefficients, and producing scatterplots with fitted lines in R, STATA, Excel, and Google Sheets
- Assignment: Using a software of your choice and following one of the online tutorials, try and open your dataset from Unit 6. Take any two variables that could plausibly be related and test your hunch in any one of the following ways:
 - generate a cross-tabulation;
 - calculate the correlation coefficient;
 - produce a scatterplot (with a fitted line), placing the variable you think is the independent variable on the x-axis and the one you think is the dependent variable on the y-axis;

Whatever test you conduct, write a summary of what you find.

• *Goal*: The main goal of this exercise is asking students to define their first bivariate hypothesis – at the end of the class, the instructor should make it clear that the students have successfully defined (and tested) their first hypothesis, which they may not have recognized at the time. In addition, they are introduced to the power (and frustration) of using statistical software to assess bivariate relationships. If at all possible, encourage students to use online guides to generate a figure or statistic that represents the bivariate association in R or another software package of their choice. Although this will be challenging for many, explain that muddling through with the software – especially through extensive internet searches – is how a lot of researchers interact with these programs, even when they have extensive experience using them.

Semester 1 Final Memo

- Assigned Readings: Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams. 2008. *The Craft of Research*. Chicago: University of Chicago Press. "From Topics to Questions." pp. 35-50, and "Using Secondary Sources to Find a Problem." pp. 88-91.
- Assignment (Coursework): Based on something we have discussed in class, something you have observed as part of your work on our team project, or something you have read or discussed elsewhere, come up with at least two puzzles i.e., research questions that you'd like to spend some time trying to answer.
- Assignment (Apprenticeship): Complete any unfinished tasks related to the team project before the end of the semester. If completing the task requires more time than you have left in the semester's apprenticeship, take detailed notes about what you have done and what is left to do, so that you or another apprentice can pick up where you left off at the start of the next semester.
- *Goal*: Give students a head-start on what is arguably the hardest part of the second semester: identifying a research question of interest to them. Separating their interest in a given question explicitly mentioned here from issues related to feasibility and integration with existing work on the topic is especially helpful. Generating two potential puzzles arms them with a spare in case one ends up not being appropriate for the second semester exercise.

A1.2 Semester 2 Coursework

Unit 1: The Evolution of the "What" and "Why" of Research

- Assigned Readings: Course syllabus and project description
- *Assignment*: In a short paragraph (and in your own words): now that you have had some experience apprenticing on our collaborative project, what do you see as its key contribution? What additional questions does the project raise that it does not currently address?
- *Goal*: Reground the team in the "why" of the project we are all collaborating on in their apprenticeship, but also emphasize the way that research (or the researcher's relationship with a project) can evolve over time.

Unit 2: Developing Research Questions

- Assigned Readings: Memo on "Identifying Good Research Questions;" Barbara Geddes. 2003. Paradigms and Sandcastles: Theory Building and Research Design in Comparative Politics. Ann Arbor: The University of Michigan Press. "Big Questions, Little Answers: How the Questions you Choose Affect the Answers You Get." pp. 27-35; Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams. 2008. The Craft of Research. Chicago: University of Chicago Press. "From Topics to Questions." pp. 35-50, and "Using Secondary Sources to Find a Problem." pp. 88-91; Matthias Lehert, Bernhard Miller and Arndt Wonka. 2007. "Increasing the Relevance of Research Questions: Consideration on Theoretical and Social Relevance in Political Science." in Gschwend and Schimmelfennig (eds.) Research Design in Political Science: How to Practice What They Preach. New York: Palgrave MacMillan. pp. 21-31, 37.
- Assignment: On your own, complete the "Identifying Good Research Questions" worksheet and come to class prepared to discuss your ideas.

- Lehert et al. identify at least eight different types of research projects that one can undertake. See if you can develop research questions of at least two different types based on your worksheet. For each, be clear about the dependent variable (outcome) of interest.
- *Goal*: The free-writing exercise is designed to help them explore their research interests in at least three areas: fascinating (or infuriating) things they learned in a classroom; topics they are repeatedly drawn to when scrolling through (social or traditional) media; and something from their lives (or the lives of their loved ones) that they wish more people understood to help make sense of a political phenomenon. Once they have identified some areas of interest, the challenge is to transform these into narrow research questions. Even though they are exposed to the variety of types of questions they could pose, in class, we emphasize how the exercises of the second semester curriculum are most successful with questions of the form "does *x* cause *y*?" The end-goal is to help every student emerge from the meeting with at least one question of this form.

Unit 3: Engaging with the Literature

- Assigned Readings: Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams. 2008. The Craft of Research. Chicago: University of Chicago Press. "Engaging Sources." pp. 84-88, 92-100.
- Additional Resources: Tips for accessing university library resources from on- and off-campus; Tom Nicholas. 2017 "How to Use Google Scholar: Finding Journal Articles and Papers Made Easy." Video; University of British Columbia iSchool. 2013. "How to Read an Academic Paper." Video.
- Assignment: Using Google Scholar or another search tool, find at least one significant article or book chapter relevant to your chosen research question. Read it strategically (spending no more than 15 minutes) and take notes on how it relates to your question.
- *Goal*: As in the first semester, the key is to convince students that they do not have to do everything on their own; neither does their research have to be entirely unique. Collaboration is critical to scientific progress. By finding existing work on their topic of interest, they need to learn how to (a) assess what has been done, by searching for relevant works, sifting through these to find something that is especially relevant, and summarizing this work effectively; and then (b) identify their contribution to it, putting their idea into conversation with the existing work.

Unit 4: Theory Development

- Assigned Readings: Paul M. Kellstedt and Guy D. Whitten. The Fundamentals of Political Science Research. Cambridge: Cambridge University Press. "The Art of Theory Building." pp. 24-46; Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams. 2008. The Craft of Research. Chicago: University of Chicago Press. "Making Good Arguments: An Overview." pp. 108-19.
- Assignment: Doing some additional background reading where necessary, try to provide a simple answer to your research question. At the very least, identify what independent variable you think explains variation in your dependent variable. Then, in a sentence or

- two, explain the mechanism linking the two. Across what dimension(s) individuals, subnational units, national units, time can you expect to see the causal relationship you have identified?
- *Goal*: Gain an understanding that, while data are useful in testing our hypotheses, we also need to be able to convince our readers that our theories are plausible even before any data is analyzed. This is often the hardest unit for many students. Encourage those who have taken game theory and/or political theory to lean on those skills to think about the logical steps linking x and y. Alternatively/additionally, help them identify these steps by (a) identifying the relevant actors involved in x and/or y; (b) thinking through the goals of these actors, especially when it comes to y; and (c) starting with x as a starting point, imagining how an actor's decisions leads into another, until y is reached.

Unit 5: Conceptualization and Measurement

- Assigned Readings: Robert Adcock and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." American Political Science Review 95 (3): pp. 530-32; Janet Buttolph Johnson, H. T. Reynolds and Jason D. Mycoff. 2016. Political Science Research Methods. Los Angeles: CQ Press. "The Building Blocks of Social Scientific Research: Measurement." pp. 121-25, 130-32, 135-62.
- Assignment: Doing additional background reading where necessary, define a measurement strategy for capturing your independent and dependent variables across at least one dimension of variation (e.g., individuals, geographic units, time). Recall that a measurement strategy should include the following steps: (i) conceptualization going from an abstract "background concept" to a well-defined "systematized concept;" (ii) operationalization defining your systematized concept in terms of one or more indicators; and (iii) a strategy for measurement, including how you would score individual cases.
- *Goal*: As in the first semester, the main goal is an appreciation for the decisions that go into conceptualizing and measuring latent variables, and that this needs to be done explicitly, and justified. Here, they develop an ideal conceptualization and measurement strategy before thinking about the practicalities of doing this with real-world data (Unit 7).

Unit 6: Ready-Made Data Sources

- *Additional Resources*: List of datasets commonly used by scholars in different subfields of political science.
- *Assignment*: Spend some time looking for existing data that capture one or more of your variables. Come to our meeting with a way of measuring at least one of
- *Goal*: As they begin to look for existing data that approximate the way they would want to measure x and y, students should gain an appreciation for the trade-offs that researchers often make between validity and feasibility. In class, help them walk through what is gained or lost by building a bespoke dataset vs. using ready-made data. Together, consider possible sources of systematic measurement error that could enter their dataset if they rely on a ready-made dataset.

Unit 7: Building a Dataset

- Assigned Readings: Julia Rathke. 2007. "Achieving Comparability of Secondary Data." in Research Design in Political Science: How to Practice What They Preach. Thomas Gschwend and Frank Schimmelfennig (eds.) New York: Palgrave MacMillan. pp. 103-24; Paul M. Kellstedt and Guy D. Whitten. 2013. The Fundamentals of Political Science Research. Cambridge: Cambridge University Press. "Getting to Know Your Data: Evaluating Measurement and Variations." pp. 109-23.
- *Additional Resources*: Links to websites that explain how to import and merge datasets, and to create and recode new variables in a number of different software packages. Additional links that explain how to generate descriptive statistics and to produce histograms and box-and-whisker plots, if available, in each software program.
- Assignment: Do what you think you need to combine or recode existing datasets to create a .csv file that includes a measure of your independent and dependent variables across your chosen units of observation. After creating your dataset, open it and describe your independent and dependent variables using any descriptive statistics you think are appropriate.
- *Goal*: Appreciation for the many steps that go into preparing data before conducting a test. Depending on the students' questions and datasets, they may need help with (a) aggregation of small units of analysis into larger ones, especially if they are using two datasets with different units of analysis; (b) merging of two or more datasets, using unique identifiers of each unit; (c) combining multiple variables to generate a measure of x or y; and/or (d) recoding variables to better capture variation in x or y.

Unit 8: Theory Testing

- Assigned Readings: Paul M. Kellstedt and Guy D. Whitten. 2013. The Fundamentals of Political Science Research. Cambridge: Cambridge University Press. "Bivariate Hypothesis Testing." and "Bivariate Regression Models." pp. 145-96.
- *Additional Resources*: Links to websites that explain how to calculate correlation coefficients, produce scatterplots (with fitted lines), and estimate linear regressions in a variety of statistical software programs.
- Assignment: Take your main hypothesis and see if you can test it in any one of the following ways: (i) create a scatterplot (with a fitted line), placing the variable you think is the independent variable on the x-axis and the one you think is the dependent variable on the y-axis; (ii) calculate the correlation between the two and see if it is statistically significant; and/or (iii) estimate whether there is a linear relationship between the two using a basic regression model. Write a summary of what you find.
- *Goal*: In addition to figuring out how to execute a test of their hypothesis using statistical software of their choice (or presenting qualitative data in an intuitive way), a critical lesson for many students is that it is perfectly acceptable to be proven wrong. Before the press "Enter," encourage them to consider how they would respond if the test fails to support their hypothesis. To do this, they may want to work backwards, reflecting on their choice of data, their conceptualization, and then their theory. What would they redo if they could? Even if

their hypothesis is supported, students are encouraged to consider whether they should adjust their measures or theory.

Semester 2 Final Memo

- Assigned Readings: Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams. 2008. *The Craft of Research.* Chicago: University of Chicago Press. "Planning." pp. 177-86.
- Assignment (Coursework): Combine all of your short assignments into a single document that includes a couple of paragraphs covering each of the following:
 - i. an introduction that motivates your choice of research question in terms of social and/or scholarly value (1 paragraph);
 - ii. a brief literature review that summarizes one or more existing theories that are relevant to your topic and concludes with a sentence or two summarizing what you see as your contribution to this existing literature (no more than 2 paragraphs);
 - iii. a definition of your hypothesis, defining your independent and dependent variables, a clear (if brief) explanation of the theoretical connection between them, and a statement of the empirical implications of this theory what would be true in the world if you were right (no more than 2 paragraphs);
 - iv. an explanation of how you define your variables and how you have attempted to measure them (1 paragraph);
 - v. descriptive statistics for each of your variables and at least one test of your theory (1 paragraph);
 - vi. a brief conclusion in which you discuss the validity of your hypothesis in light of your test(s) and consider next steps (no more than 2 paragraphs).

Altogether, this document need not exceed 5 pages.

- Assignment (Apprenticeship): Complete any unfinished tasks related to the team project before the end of the semester. If completing the task requires more time than you have left in the semester's apprenticeship, take detailed notes about what you have done and what is left to do, so that you or another apprentice can pick up where you left off at the start of the next semester/year.
- *Goal*: Putting together a (small) research paper from start to finish. Emphasize that importance of (iv) above, especially when it comes to reflecting in what could have been different working backwards from the choice of data, their conceptualization, and then their theory. What would they redo if they could? Regardless, what would their next steps be to adjust or expand upon the project?

A1.3 Semester 3 Coursework

Unit 1: The Structure of Research

• Assigned Readings: Memo on "The Structure of Research;" Tom Nicholas. 2017 "How to Use Google Scholar: Finding Journal Articles and Papers Made Easy." Video; University of British Columbia iSchool. 2013. "How to Read an Academic Paper." Video.

- Assignment (Coursework): Start exploring broad topics of interest to you by using Google Scholar, reading strategically to cover as many readings as possible.
- Assignment (Apprenticeship): You are now senior apprentices, meaning you are in a great position to help get the new, junior apprentices ready to complete tasks for our team project. Be generous with your time and encourage them to come to you with questions, especially as they work through their training exercises. If there are tips that you think would be helpful for the whole group, feel free to post these to our team discussion channels. If there are questions that you feel you cannot answer, see if you can team up with another senior intern or ask me for help.
- *Goal*: Prompt the students to consider areas of interest to them for this larger, more complex research project. Help them recognize whether they want to continue working on their project from the previous semester or not. Either way, they should use their experience to inform their next steps, including lessons learned about what works and what not to do. Remind them that a project related to their apprenticeship could result in a co-authored article, but emphasize that there is no pressure to deviate from their personal interests.

Unit 2: Identifying a Question

- Assigned Readings: Memo on "Identifying Good Research Questions;" Barbara Geddes. 2003. Paradigms and Sandcastles: Theory Building and Research Design in Comparative Politics. Chapter 2 ("Big Questions, Little Answers: How the Questions you Choose Affect the Answers You Get"), pp. 27-35; Matthias Lehert, Bernhard Miller and Arndt Wonka. 2007. "Designing Research in Political Science: A Dialogue between Theory and Data." in Gschwend and Schimmelfennig (eds.) Research Design in Political Science: How to Practice What They Preach, pp. 21-31.
- Assignment: Identify at least two research questions that you might be interested in trying to answer. For each, identify the dependent variable.
- *Goal*: As with the last semester, asking students to develop two research questions helps ensure that they will leave the meeting with at least one feasible idea for a project. They will almost certainly be more drawn to one over the either, and do whatever possible to help them transform that question into something doable over the course of two semesters. Emphasize that a good project must be be interesting to the researcher ("personal significance") and interesting to other researchers ("scientific significance"). It could also potentially be interesting to the wider world ("social significance"). Critically for their purposes, the project must be feasible given the timeframe, their resources, and skills. Make clear that you can help them develop necessary skills if they want to aim to expand upon the project later.

Unit 3: Developing Good Theories

• Assigned Readings:: Barbara Geddes. 2003. Paradigms and Sandcastles: Theory Building and Research Design in Comparative Politics. Chapter 5 ("How the Approach You Choose Affects the Answers You Get: Rational Choice and its Uses in Comparative Politics"), pp. 175-211; Kenneth Hoover and Todd Donovan. 2011. The Elements of Social Scientific Thinking, 10th Edition. Boston: Wadsworth. Chapter 4 ("Refinements"), pp. 58-77.

- *Assignment*: Take one or two of your proposed research questions and think through a possible answer (or two). What is the logic connecting the question and your answer?
- *Goal*: As in the second semester, the goal is to convince themselves and their readers that their hypothesis is sound. As before, they will want to define the logical steps linking x and y by (a) identifying the relevant actors involved in x and/or y; (b) thinking through the goals of these actors, especially when it comes to y; and (c) starting with x as a starting point, imagining how an actor's decisions leads into another, until y is reached. Drawing a flowchart or game tree may be helpful.

Unit 4: Developing Good Tests

- Assigned Readings: Kenneth Hoover and Todd Donovan. 2011. The Elements of Social Scientific Thinking, 10th Edition. Boston: Wadsworth. Chapter 5 ("Measuring Variables and Relationships"), pp. 78-117; Joshua D. Angrist and Jörn-Steffen Pischke. 2008. Mostly Harmless Econometrics: An Empiricist's Companion. Chapter 2 ("The Experimental Ideal"), pp. 9-18.
- *Assignment*: Take one or two of your proposed question/answer combinations and think through how you would go about testing it.
- *Goal*: As before, we want to consider the ideal test of a theory before getting caught up in the practicalities of executing it. At a minimum, students should define their conceptualization and measurement strategies and consider what type of comparison(s) would be the best test of their hypothesis.

Unit 5: Creating Your Dataset

- Assigned Readings: Janet Buttolph Johnson, H. T. Reynolds and Jason D. Mycoff. 2016. Political Science Research Methods. Los Angeles: CQ Press. "The Precision of Measurement." pp. 150-62; Julia Rathke. 2007. "Achieving Comparability of Secondary Data." in Gschwend and Schimmelfennig (eds.) Research Design in Political Science: How to Practice What They Preach, pp. 103-24; Paul M. Kellstedt and Guy D. Whitten. 2013. The Fundamentals of Political Science Research. Cambridge: Cambridge University Press. "Getting to Know Your Data Statistically," "What is the Variable's Measurement Metric," "Describing Categorical Variables," "Describing Continiuous Variables," and "Limitations of Descriptive Statistics and Graphs." pp. 109-23.
- Additional Resources: List of datasets commonly used by scholars in different subfields of political science. Links to websites that explain how to import and merge datasets, and to create and recode new variables in different software packages. Additional links that explain how to generate descriptive statistics and to produce histograms and box-and-whisker plots, if available, in each software program.
- Assignment: Pick one of your questions and one of your tests that you think will be the focus on your paper. Spend some time looking for datasets that might be a good way of capturing the dependent and independent variables you have identified in your hypothesis. At what unit of analysis are they available? Do you need to recode the data in any way? Describe your choice of dataset in a paragraph or two, explaining what (if any) recoding

you needed to do. Then take some time to get to know your data. Are there any patterns in each individual variable that you can already identify? Do they seem to operate as you would expect? You may want to create some graphs or tables that illustrate what you've found and write up a paragraph or two describing them. (This will comprise part of the "Data" section of your paper.)

• *Goal*: Again, understand the trade-off between the ideal and the practical. This is especially true with the students who want to use survey data to test their theories: they can either choose to make use of (less valid) existing datasets or design their own survey. In semesters when multiple students design their own surveys, theses can be combined into an omnibus survey and distributed among a convenience sample by the students and the PI. For students using ready-made data, they will need to consider whether datasets need to be merged, and whether any variables need to be aggregated, combined, and/or recoded. Small group discussion, based on type of data used, can be helpful at this stage.

Unit 6: Initial Testing

- *Additional Resources*: Links to websites that explain how to calculate correlation coefficients, produce scatterplots (with fitted lines), and estimate linear regressions in a variety of statistical software programs.
- Assignment: Conduct some initial tests of your theory and create a couple of visualizations of the relationship between the independent and dependent variables you have identified. Do your initial findings seem to support your hypothesis? Come to our meeting with one or two illustrations of the test you conducted, ready to discuss how you think they fit into your broader argument.
- *Goal*: Again, figure out how to conduct a test in their statistical software of choice or present qualitative evidence in an intuitive way. Help students think about some key questions in data visualization by asking them to present one figure to the group, with other students providing feedback on what is not (yet) clear. But many students will not have data at this point, especially those creating bespoke datasets. They should work on defining the steps that they will take to conduct their test once the dataset is complete, including the comparison(s) they will make and the evidence they will look for to assess their hypothesis.

Unit 7: Discussing Results

- Assigned Readings: Dirk De Bièvre. 2007. "Falsification in Theory-Guided Empirical Social Research: How to Change a Tire while Riding Your Bicycle." in Gschwend and Schimmelfennig (eds.) Research Design in Political Science: How to Practice What They Preach, pp. 203-15.
- *Assignment*: Write up a couple of paragraphs describing the results from Unit 6. These will comprise the "Discussion" section of your draft.
- *Goal*: Teaching students how to be reflective of test results whether confirming their hypotheses or otherwise is key. If the results seem to support their hypothesis, can they interrogate the test to confirm the robustness of the results? If the results seem to disprove their hypothesis, what are the next steps? To understand what should come next, students

will want to work backwards, reflecting on their choice of data, their conceptualization, and then their theory. What should they redo? These next steps should be part of the "Discussion" section of their first draft.

Completing a First Draft (Semester 3)

- Assignment (Coursework): Put together the different pieces you have been working on
 over the past weeks and try to create a seamless first draft of your paper. Don't forget to
 motivate your paper with a compelling introduction and to conclude with some avenues for
 future research.
- Assignment (Apprenticeship): Complete any unfinished tasks related to the team project before the end of the semester. If completing the task requires more time than you have left in the semester's apprenticeship, take detailed notes about what you have done and what is left to do, so that you or another apprentice can pick up where you left off at the start of the next semester.

A1.4 Semester 4 Coursework

Unit 1: Re-Introduction to Research Projects

- *Assignment*: Re-familiarize yourself with your research project. Come to our meeting prepared to discuss the following:
 - i. Your research question and hypothesis.
 - ii. Any (preliminary) data you have collected and analyzed.
 - iii. Your next steps, whether (additional) data collection and/or analysis, reformulating your question and/or hypothesis, or something else.

Unit 2: Re-Defining and Re-Framing Your Question

- Assigned Readings: Refer back to readings from Semester 3, Unit 2.
- Assignment: Submit a draft introduction for your research paper (approximately 1-2 pages). Be sure to motivate your project, define your question as well as what you see as your contribution, while providing a brief roadmap of the paper you intend to write.

Unit 3: Building or Expanding a Literature Review

- Assigned Readings: Refer back to readings from Semester 2, Unit 3.
- Assignment: A key aspect of a good research project is building upon and contributing to an existing literature, making sure that you are doing something that is new (but not too new). Using Google scholar and library resources, identify some if the literature(s) relevant to your paper. Read them strategically and write a draft literature review that identifies:
 - i. The state of the existing literature; and ii. What you think you are contributing to it. This should be accomplished in approximately 3-5 pages.

Unit 4: Re-Developing Your Theory

• Assigned Readings: Refer back to readings from Semester 3, Unit 3.

- Assignment: Your theory section should make clear what you see as your answer to your research question. More than likely, your theory is based one or more assumptions that you are making about the world as you see it. Your draft theory section should make clear the argument you are making by:
 - i. Taking your readers through the logical steps you are making, connecting changes in your independent variable(s) to a change in your dependent variable; and
 - ii Being explicit about what assumptions you are making. In both cases, you may rest on arguments made by other scholars. You might also consider some alternative arguments that could be or have been made and why you disagree with them. Your theory section should end with one or more hypotheses that make explicit what we would see in the world if your theory (and not any of the alternative ones) were correct.

This should be accomplished in approximately 3-5 pages.

Unit 5: Re-Creating and Describing Your Data

- Assigned Readings: Refer back to readings from Semester 3, Unit 5.
- Assignment: A section of your paper should describe the data you are using to test your theory, whether these data are quantitative or qualitative. In addition to describing the source of your data, describe how you are measuring each of your variables, including whether you are combining or recoding any existing data. Consider how this measurement strategy differs from an ideal one, discussing whether you are concerned about any systematic measurement error that could bias your results. You might also consider including some descriptive statistics. Depending on the complexity of your data, the section should comprise 3-5 pages.

Unit 6: Re-Testing Your Theory

- Assigned Readings: Refer back to readings from Semester 3, Unit 6.
- Assignment: Conduct some initial tests of your theory and create a couple of visualizations of the relationship between the independent and dependent variables you have identified. Describe what you see and consider whether these initial findings seem to support your hypotheses (approximately 3-5 pages, including figures).

Unit 7: Re-Discussing Your Results

- Assigned Readings: Refer back to readings from Semester 3, Unit 7.
- Assignment: Write up a couple of paragraphs describing your results and how they relate to your theory. Are there reasons to be skeptical of the patterns you see, either because of the data you are using or the test you are conducting? Are there additional tests that should now be conducted? Your discussion section should comprise 2-3 pages.

Completing a Paper (Semester 4)

• Assignment (Coursework):: Put together the different pieces you have been working on over the past weeks and try to create a seamless presentation of your project. Don't forget

- to motivate your paper with a compelling introduction and to conclude with some avenues for future research.
- Assignment (Apprenticeship): Complete any unfinished tasks related to the team project before the end of the semester. If completing the task requires more time than you have left in the semester's apprenticeship, take detailed notes about what you have done and what is left to do, so that you or another apprentice can pick up where you left off at the start of the next semester/year.

A2 Survey of Current and Former Students

A2.1 Method

61 former and current program participants were invited to participate in an online survey, hosted on Qualtrics in December 2022. After seven days, the survey closed, at which point 33 (54.1%) had completed the survey. Participant recruitment, survey instrument, and data management were all approved by the Institutional Review Board of a large R1 university.

Importantly, no identifying information was explicitly collected, and the recruitment materials emphasized that participation was voluntary and that their responses would be protected in the following ways: (i) any identifying information in open-ended responses would be redacted prior to analysis and distribution; (ii) analysis of multiple-choice responses would be done across the full set of responses, so that no response is identifiable; (iii) analysis by sub-groups would be done only if there were at least two (2) respondents in each sub-group; and (iv) no uniquely identifiable information would accompany quotes from open-ended responses. Participants were also assured that nothing they share would ever influence the way they are evaluated, including in letters of recommendation. Prior to taking the survey, participants were presented with a consent form which reminded them of these points and shared information about the purpose of the study. Before beginning the survey, they were asked to confirm that they had read and understood the consent form. All 33 participants indicated that they read and understood the form.

A2.2 Survey Instrument and Descriptive Statistics

The following questions were asked of all participants, with descriptive statistics for all multiple-choice questions. All questions appear except those who are unique to the specific program and would undermine the blind review process. Questions are offered verbatim, except where information is specific to the program or the institution. In these cases, some text has been changed or redacted with text inside of square brackets.

What year did you graduate or do you expect to graduate from [the university]?

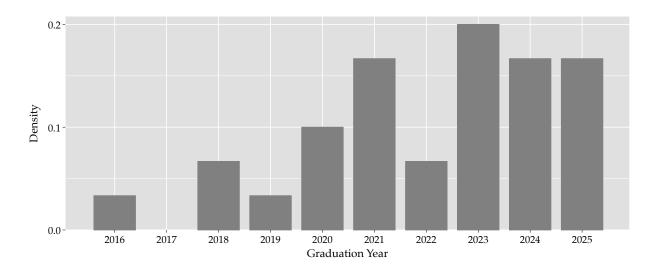


Figure A1: Respondent Graduation Year

Note: In some analyses, students who participated before multiplesemesters were offered (N = 4) and those who recently completed their first semester in the program and have not yet had the opportunity to participate in multiple semesters (N = 8) are excluded from the analysis..

While at [the university], were/are you a Political Science major or minor?

- Yes, it was my main/only major. (72.7%)
- Yes, it was my secondary major. (21.21%)
- Yes, it was my minor. (6.1%)
- No. (0%)

If Political Science was/is not your (only) major, what (else) was/is your major? 18 responses: Economics (33.3%); History (16.7%); Geography (16.7%); Spanish (11.1%)

How many semesters did/have you participate(d) in [the] program?

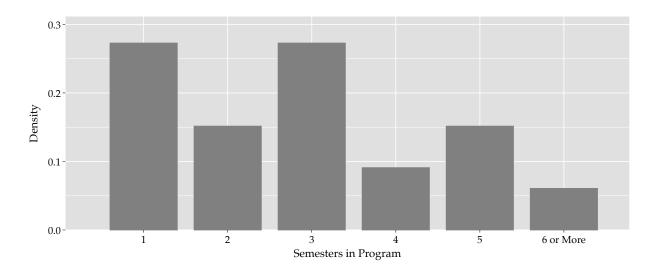


Figure A2: Semesters in Program

The next section of questions focuses on the time **before you started [the] program**. To the best of your ability, think back to that time and answer the questions as best as you can. Before participating in [the] program...how much did you know about **social science research**?

- A lot. (6.1%)
- A little. (39.4%)
- Not much. (33.3%)
- Nothing. (21.2%)

Before participating...how **interested** were you in social science research?

- Very interested. (30.3%)
- Somewhat interested. (51.5%)
- Not very interested. (18.2%)
- Not at all interested. (0%)

Before participating...did you have plans to complete a **senior thesis** or capstone project?

- Yes, I had already applied to the honors program. (9.1%)
- I was planning to. (15.2%)
- I was considering it. (30.3%)
- I had decided against it. (6.1%)
- I hadn't really thought about it. (39.4%)

Before participating...when you thought about your **plans for after graduation**, did these involve research in any way?

• Yes, I was planning to pursue a post-graduate program/job that would involve research. (5.0%)

- Yes, I was considering a post-graduate program/job that would involve research. (39.4%)
- No, I had planned to pursue a post-graduate program/job that wouldn't involve research. (36.4%)
- I hadn't really thought about it. (9.1%)

The next set of questions involve the **[apprenticeship] component** of [the] program. By this, I mean the weekly hours that you spent working on our group project []. (Soon, I'll ask you some questions about the *coursework* component, including your independent research projects.)

How much do/did you **enjoy** [the apprenticeship]?

- Really enjoy(ed) it. (27.3%)
- Enjoy(ed) it. (66.7%)
- Don't/didn't really enjoy it. (6.1%)
- Don't/didn't enjoy it at all. (0%)

How much did you learn **about the research process** from your [apprenticeship]?

- I learned a lot. (63.6%)
- I learned a little. (33.3%)
- I didn't learn much. (3.0%)
- I didn't learn anything. (0.0%)

Do you think the [apprenticeship] helped you gain professional skills?

- Yes, I gained many skills. (39.4%)
- Yes, I gained some skills. (57.6%)
- No, I didn't gain many skills. (3.0%)
- No, I didn't gain any skills. (0.0%)

If you gained any professional skills, which skill(s) do you believe you developed through the [apprenticeship]? (Check all that apply)

- Written communication:
 - ...with my peers.
 - ...with more-junior team-members.
 - ...with more-senior team-members, including with my supervisor.
- Verbal communication:
 - ...with my peers.
 - ...with more-junior team-members.
 - ...with more-senior team-members, including my supervisor.
- Organizational skills
- Punctuality
- Other (specify)

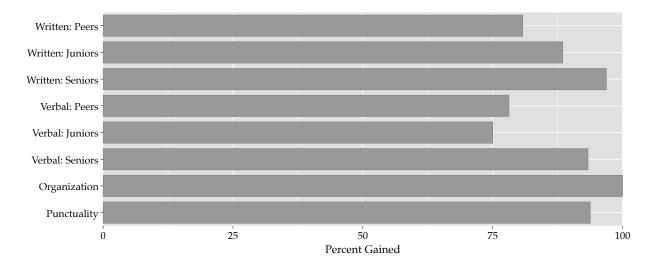


Figure A3: Skills Gained: Apprenticeship

Thinking back on your [apprenticeship], are there particular experiences that were **particularly impactful**? If so, can you share these?

Thinking back on your [apprenticeship], is there anything you wish could have been **different** about the project or your [apprenticeship] experience?

Is there anything else you'd like to share about your experience with the **[apprenticeship] component** of [the] program?

This set of questions involve the **coursework component** of [the] program. By this, I mean the readings and assignments you completed according to the different syllabi, the independent research projects you developed as part of the program, and the meetings we had to discuss any of these.

How would you evaluate the coursework in terms of...teaching you about the **purpose of social science research**?

- Effective I learned a lot. (54.8%)
- Somewhat effective I learned some things. (41.9%)
- Not very effective I didn't learn much. (3.2%)
- Not at all effective I didn't learn anything. (0%)

How would you evaluate the coursework in terms of...helping you <u>learn</u> the steps of the research process?

- Effective I learned a lot. (67.7%)
- Somewhat effective I learned some things. (32.3%)
- Not very effective I didn't learn much. (0%)
- Not at all effective I didn't learn anything. (0%)

How would you evaluate the coursework in terms of...helping you <u>apply</u> the steps of the research process to your own research interests?

- Effective I learned a lot. (61.3%)
- Somewhat effective I learned some things. (35.4%)
- Not very effective I didn't learn much. (3.2%)
- Not at all effective I didn't learn anything. (0%)

How would you evaluate the coursework in terms of...helping you **develop your own research project/interests**?

- Effective I learned a lot. (61.3%)
- Somewhat effective I learned some things. (32.3%)
- Not very effective I didn't learn much. (6.5%)
- Not at all effective I didn't learn anything. (0%)

How would you evaluate the coursework in terms of...helping you succeed in the [apprenticeship] component of [the] program?

- Effective it helped a lot. (35.5%)
- Somewhat effective it helped somewhat. (35.5%)
- Not very effective it didn't help much. (29.0%)
- Not at all effective it didn't help at all. (0%)

Do you think the coursework component of the program helped you in your **other classes**?

- Yes, I gained many skills that I used in many of my other classes. (40.0%)
- Yes, I gained some skills that I used in some of my other classes. (60.0%)
- No, I didn't gain many skills that were helpful in my other classes. (0%)
- No, I didn't gain any skills. (0%)

If you gained any skills that were helpful in other classes, which skill(s) do you believe you developed through our coursework? (Check all that apply)

- Reading social science research.
- Identifying existing literature relevant to a topic.
- Strategically reading the literature on a topic.
- Writing a review of the existing literature.
- Identifying an interesting research question.
- Developing a theory.
- Deciding how variables should be defined.
- Deciding how variables should be measured.
- Identifying existing data sources.
- Evaluating existing data sources.
- Developing new datasets, including recoding existing data.
- Basic data analysis.
- Discussing test results.

- Writing a complete research paper.
- Other (specify)

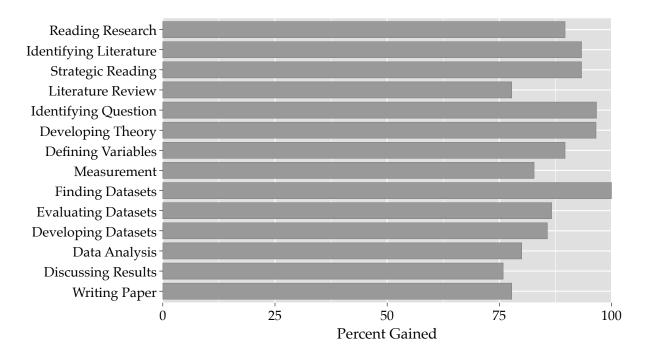


Figure A4: Skills Gained: Coursework

Thinking back on the coursework that you completed, are there particular lessons that were **particularly impactful**? If so, can you share these?

Thinking back on the coursework that you completed, is there anything you wish could have been **different** about the course?

Is there anything else you'd like to share about your experience with the **coursework component** of [the] program?

Did you write – or are you planning on writing – a senior thesis or other capstone project?

- Yes, I wrote or am currently writing a thesis or other capstone project. (42.4%)
- Yes, I am planning on writing a thesis or other capstone project. (21.2%)
- No, I didn't write (or have any plans on writing) a thesis. (36.4%)

If you wrote or are writing a thesis/are planning on writing a thesis...did your experience with our program **influence your decision**?

- Yes, it (has) made me much more likely to write a senior thesis or capstone project. (47.6%)
- Yes, it (has) made me somewhat more likely to write a senior thesis or capstone project. (38.1%)

- Yes, it (has) made me less likely to write a senior thesis or capstone project. (0%)
- No, it didn't have/hasn't had any effect. (14.3%)

If you didn't write (or don't have plans to write) a thesis...how did your experience with our program **influence your decision**?

- It made me more likely to write a thesis, even though I still decided not to do it. (54.6%)
- It had no effect, one way or the other. (36.4%)
- It made me somewhat less likely to write a thesis. (9.1%)
- It made me decide not to write a thesis. (0%)

If you wrote or are writing a thesis...how helpful was the **[apprenticeship] component** of [the] program in preparing you to write a successful thesis?

If you are planning on writing a thesis...how helpful do you think the **[apprenticeship] component** of [the] program will be in preparing you to write a successful thesis?

- Very helpful. (61.9%)
- Somewhat helpful. (33.3%)
- Not particularly helpful. (0%)
- Not at all helpful. (4.8%)

If you wrote or are writing a thesis...how helpful was the **coursework component** of [the] program in preparing you to write a successful thesis?

If you are planning on writing a thesis...how helpful do you think the **coursework component** of [the] program will be in preparing you to write a successful thesis?

- Very helpful. (65%)
- Somewhat helpful. (35%)
- Not particularly helpful. (0%)
- Not at all helpful. (0%)

Is there anything else you'd like to share about the role [the] program played in your [thesiswriting] process?

After beginning [the] program, did you ever work as a **research assistant (RA)** for me or another faculty member? (To clarify, RAs on my team are paid hourly; RAing for other faculty members could involve paid work or working for course credit...but without any coursework component.)

- Yes, for you. (36.4%)
- Yes, for someone else. (15.2%)
- No/not yet. (51.5%)

If you worked as an RA...was the position paid, done for credit, or something else?

- Paid. (39.4%)
- Unpaid, for course credit. (6.1%)

- Unpaid, with no compensation. (3.0%)
- Other (specify)

If you worked as an RA...how helpful was the **[apprenticeship] component** of [the] program in preparing you for this work?

- Very helpful. (50%)
- Somewhat helpful. (50%)
- Not very helpful. (0%)
- Not at all helpful. (0%)

If you worked as an RA...how helpful was the **coursework component** of [the] program in preparing you for this work?

- Very helpful. (35.7%)
- Somewhat helpful. (42.9%)
- Not very helpful. (21.4%)
- Not at all helpful. (0%)

If you worked as an RA...is there anything else you'd like to add about the role [the] program played in this process?

If you didn't work/haven't worked as an RA...is there anything you'd like to add about the role [the] program played in this process?

What are you up to these days?

- I'm still an undergraduate at [the university]. (51.5%)
- I'm in a post-graduate program at [the university] or another university. (18.2%)
- I'm looking for work. (9.1%)
- I'm currently working. (21.2%)
- Other (please specify)

For those who have graduated:

- Can you say more about your current position/program?
- Did you do anything between graduating from UIUC and your current position/program? If so, what did you do?
- Is there anything you'd like to share about the role of [the] program in your journey since graduation?

For those who have not yet graduated:

• Have you given any thought to what you might do after graduation? If so, what are you considering doing?

• Is there anything you'd like to share about the role of [the] program in your post-graduation plans?

Open-ended responses about current and future post-graduate plans were hand-coded as involving research (1) or not (0) by the PI. (In future iterations, survey participants should be asked to evaluate this for themselves, to avoid measurement error.)

- Related to research. (57.1%)
- Unrelated to research. (42.9%)

Finally, just a few more questions about you and your identity.

How old are you? μ =22.2, σ =2.44 [19, 28]

In terms of your gender, how do you identify?

- Cis-gender male (42.4%)
- Cis-gender female (57.6%)
- Transgender male (0%)
- Transgender female (0%)
- Non-binary (0%)
- Prefer not to say (0%)
- Other (0%)

Do you consider yourself...a member of the LGBTQ+ community? (Feel free to add more information about your identity/identities in the space provided.)

- Yes (30.3%)
- No (63.6%)
- Other (3.0%)
- Prefer not to say (3.0%)

Do you consider yourself...a member of a historically marginalized racial group? (Feel free to add more information about your identity/identities in the space provided.)

- Yes (30.3%)
- No (57.6%)
- Other (6.1%)
- Prefer not to say (6.1%)

Do you consider yourself...a member of a historically marginalized ethnic group? (Feel free to add more information about your identity/identities in the space provided.)

- Yes (37.5%)
- No (56.3%)
- Other (0%)
- Prefer not to say (6.3%)

Do you identify as "first-gen," i.e., are you the first member of your immediate family to attend university? (Feel free to add more information about your identity/identities in the space provided.)

- Yes (24.2%)
- No (75.8%)
- Other (0%)
- Prefer not to say (0%)

Do you consider yourself...a person with a disability? (Feel free to add more information about your identity/identities in the space provided.)

- Yes (3.0%)
- No (90.9%)
- Other (3.0%)
- Prefer not to say (3.0%)

Do you consider yourself...a veteran of the armed forces? (Feel free to add more information about your identity/identities in the space provided.)

- Yes (0%)
- No (100%)
- Other (0%)
- Prefer not to say (0%)

Thank you for taking the time to fill out this survey. I know that the program will be better for your feedback. As always, if you have any questions or concerns, I hope you won't hesitate to reach out to me.

A3 Data on Participant and Non-Participant Performance

A3.1 Data Collection

Anonymized demographic information, enrollments, and grades for all political science majors between Fall 2013 and Summer 2021 at a large R1 university (N = 1305) were collected and sent to the author, in accordance with a research protocol approved by the university's Institutional Review Board. All political science majors who started their university degree in Fall 2013 or later were included, including those who began their training at another institution and transferred to the university before graduation, as well as those who began their degrees undeclared or majoring in another field, but who eventually declared or added a major in political science. Those who started as political science majors but switched to another program before graduation were excluded. All students who appeared in the dataset graduated in or before Summer 2021.

The original data request also included all political science minors, but as this was a much larger group, the university's data collection team was not able to gather the information within a reasonable timeframe. The survey of former and current participants (see Appendix A2) indicates that all were political science majors (even if it was not their primary major), so students with a

political science minor are unlikely to be an appropriate "control" group to study the effect of the program's "treatment."

A3.2 Measures and Descriptive Statistics

Two types of data were requested: demographic information and information about student performance in each semester. Below, I offer descriptive statistics for both participants in the program and for non-participants. Program participants were identified based on their enrollment in one or more of the courses that accompany the program. Because students can enroll in these courses with other faculty members, unique course reference numbers were used to identify the courses led by the author. Using this method, 25 program participants were identified, missing only 2 of 27 program participants who fit within these parameters. (The missing participants enrolled in similar courses based in different departments and so were not able to be identified given available data.

The number of courses completed was used as an indicator of the number of semesters the students remained in the program (2.44 on average).

Available **demographics** included the following:

- Students' **gender**, whether male (52.0% of participants; 52.3% of non-participants) or female (48.0%; 47.7%).
- Their **race** or **ethnicity**, whether White (44.0% of participants; 59.9% of non-participants), Hispanic (12.0%; 17.1%), Black (8.0%; 7.1%), Asian (24.0%; 8.2%), Mixed (4.0%; 4.8%) or Other (8.0%; 3.0%).
- Whether they began their degree at the university (96.0% of participants; 76.9% of non-participants) or **transferred** in (4.0%; 23.1%).

Student achievement was measured using the following data:

- Students' cumulative **grade point average (GPA)** was available at each semester. GPAs ranged from 2.95 to 3.99 (μ = 3.71, σ = 0.30) for participants and 1.91 to 4.0 (μ = 3.40, σ = 0.42) for non-participants.
- The average **course difficulty** was calculated using the numbers of courses that students completed. In this political science department, there is no distinction between courses at the 100- and 200-level and those at the 300- and 400-level, so the share of advanced (300- or 400-level) political science courses was calculated for each student in each semester, using all political science courses as the denominator and excluding courses in other departments. The share of advanced courses was 4.1% for participants and 4.8% for non-participants in their first year at the university, and 34.7% compared to 30.2% over their full university careers.
- Completion of a **senior thesis** was identified using the specific courses associated with the thesis program in the department. Senior thesis or capstone projects completed in other departments were therefore not included. The rate of thesis completion among program participants was 36.0%, compared to 7.3% among non-participants.

A3.3 Methods

Participants were matched with at least one similar non-participants using both demographic data and pre-trend achievement data. Nearest neighbors were identified for each participant, requiring an exact match on transfer status, matriculation and graduation year, gender, and race, as well as proximity on GPA and course difficulty in the students' first year, i.e., before they were eligible for participation in the program. Participants were then compared with their match(es) on their final cumulative GPA, their average course difficulty across all years, and whether or not they completed a senior thesis. The average treatment of participation was calculated using logit models, and all three were statistically significant at conventional levels:

	Coefficient	Robust Std Error	P > z
Final GPA	0.115	0.067	0.085
Advanced Coursework	0.035	0.006	0.000
Senior Thesis	0.340	0.123	0.006

Additional analysis compared participants who had participated in the program over multiple semesters (N = 16) and those who had completed just one semester in the program (N = 25). These two groups were compared across the same three outcomes using two-sample t-tests. The results were as follows, with means and standard errors reported for each group:

	Single Semester	Multiple Semesters	Pr (T > t)
Final GPA	3.61 (0.128)	3.76 (0.058)	0.236
Advanced Coursework	0.313 (0.029)	0.366 (0.013)	0.073
Senior Thesis	0.222 (0.147)	0.438 (0.127)	0.301