Intermediate Methodology
POLI:5003
Department of Political Science
College of Liberal Arts and Sciences
University of Iowa
Spring 2026

Instructor: Prof. Colin Case
Meeting Time: Tu 2-4:50 & Th 12:30-1:20
Classroom: 103 Schaeffer Hall (SH)
Email: colin-case@uiowa.edu
Office: 359 Schaeffer Hall
Office Hours: Tu 11-12:30 & Th 1:30-3:00
DEO: Julie Pachecco
DEO Email: julianna-pacheco@uiowa.edu

COURSE DESCRIPTION

This course is an introduction to the theory and applied use of linear regression for political scientists. It is designed to familiarize you with basic concepts in regression with Ordinary Least Squares (OLS) as the core framework. We will begin the semester by covering its assumptions, basic properties, and diagnostic tools for violations of these assumptions. We will also explore the use and interpretation of continuous, ordinal, nominal, and indicator independent variables, as well as their interactions. After covering these topics, we will move on to a conceptual introduction to causal inference, with a specific focus on when and where OLS estimates are causally identified. We will end the semester by discussing several design-based tools for causal identification that address the limitations of "vanilla" OLS in observational research.

In addition to covering the foundations of OLS, a portion of this course will also be dedicated to learning and developing proficiency with the computing and statistical environment R. R is a free, open-source statistical software that will enable you to carry out both basic and advanced statistical methods, which you will see in advanced methods courses such as networks or text analysis. In this course, we will primarily focus on coding principles to develop a strong baseline proficiency and familiarity with R that will allow you to continually learn new tools and skills as your research interests develop. There are other statistical software programs you may use and will likely encounter in graduate school (Stata, Python, etc.). Even if you do not use R long-term to conduct statistical analysis, the intuition you develop from this language will benefit you in learning other statistical languages and writing better code.

This course is central to your training as a quantitative social scientist and is likely the most important in the methods sequence. You already have a basic understanding of probability and probability distributions, hypothesis testing, and some simple statistics. You have also been introduced to questions of measurement, concepts like reliability and validity, and many aspects of research design. In short, you have many of the basic building blocks in place to begin doing real, systematic, quantitative social science research, but you likely lack the tools to apply them systematically. This course will help you become an increasingly skilled user of statistical methods and a critical consumer of research that employs them.

You will not be experts by the end of this course (or even close to it). However, you will have built a very strong foundation for future coursework, you will be better positioned to read more of the literature in your field, and you will be well-versed in using OLS regression to answer your own research questions. OLS and the linear model provide the basis for most published empirical political science research, and the jumping-off point for most of the more sophisticated methods you will see and use. As a result, this course is extremely important in your Ph.D. training. I strongly encourage you to push yourself to get as much out of this course as you possibly can.

LEARNING OBJECTIVES

By the end of the course, you will be able to…

- Understand the fundamental logic of Ordinary Least Squares (OLS) regression, including its core assumptions, the interpretation of model components and estimates, and the consequences of violating those assumptions.
- Apply the principles of OLS regression to novel empirical settings by identifying appropriate model specifications, diagnosing potential violations of assumptions, and assessing how those issues affect interpretation and inference.
- Understand the conditions under which OLS estimates can be interpreted causally, including the role of research design, identification assumptions, and threats to causal inference.
- Apply causal reasoning to empirical research by evaluating whether regression-based estimates are plausibly causal and by selecting appropriate design-based strategies to address identification problems.
- Understand the basic structure and logic of R as a statistical programming environment, including core data objects, syntax, and how statistical methods are implemented in code.
- Apply fundamental R programming skills to manipulate data and carry out regression-based analyses using real political science datasets.

CLASS STRUCTURE AND EXPECTATIONS

This is a graduate methods course. As such, I expect you to be professional, engaged, and committed to your growth and development as a researcher. This means coming to class and arriving on time, completing assignments before due dates, and behaving professionally and collegially with myself and others in the class. I also expect that you will have done the readings and assignments on time and before class starts.

The course will have two meeting times during the week. Our Tuesday meeting will be a 3-hour lecture focusing on statistical and mathematical concepts related to the course. During this session, laptops are not permitted, and you must take notes by hand (a tablet is permitted as long as notes are handwritten and not typed) unless you have an approved accommodation. I will be using both slides and a chalkboard during these sessions. Slides will typically be posted to ICON the night before class, so you can print them out if you would like.

Our Thursday meeting will be a 1-hour lab session focused on working in the R programming language. You should bring a laptop that supports R to these sessions. In these sessions, we will be live coding through an exercise. Before class, I will post (1) an unfilled class exercise, (2) a filled class exercise, and (3) any relevant data. You are more than welcome to use these scripts as you see fit to aid your learning best. I recommend following along and coding in the unfilled script with the filled script up, so if you miss a command during the class session, you can get it from the filled script. The materials will be posted before class, and I expect you to have downloaded them beforehand and installed any packages relevant to the class so we can start right at the beginning of class. This does not mean you should be working through exercises before class starts.

As a methods course, what we will be doing throughout the semester is inherently mathematical at some level. However, it is important to remember that the goal of working through mathematical derivations is to develop a deeper understanding of the concepts being covered. It is developing this understanding, not the math itself, that is the course's primary goal. At times, I will present statistical concepts in a more "intuitive" way to help you better understand them. That intuition is not a substitute for the math, nor do I mean that you should be satisfied with some general sense of what is going on without understanding the math. What I do mean is that understanding the logic of quantitative analysis runs deeper than just a set of mathematical rules and formulas. A constant theme in the course will be why a practicing political scientist would want to know about the statistical topic at hand. If I fail to make it clear at any point in the semester why we are learning what we are, you should make sure I do.

REQUIRED TEXTS

There is a required textbook for this course that all students must have access to:

> Jeffrey M. Wooldridge, Introductory Econometrics: A Modern Approach, 8th ed.

Students have the option of either purchasing the textbook themselves or using the ICON Direct version (more cost-efficient). Students considering taking qualifying exams in methods are encouraged to purchase a physical copy of the textbook for reference when studying in the future. Additional readings will be linked on the calendar and freely available online.

COURSE REQUIREMENTS AND GRADING

Grading Ranges

A+: 100-99.00
A: 98.99-93.00
A-: 92.99-90.00
B+: 89.99-87.00
B: 86.99-83.00
B-: 82.99-80.00
C+: 79.99-77.00
C: 76.99-73.00
C-: 72.99-70.00
D+: 69.99-67.00
D: 66.99-63.00
D-: 62.99-60.00
F: 59.99-00.00

Course Components

<u>Midterm and Final (25% each; 50% total):</u> The class will have two exams, a midterm and a final. The midterm exam will take place on March 10th and will be held in class. You will have 2 hours for the exam, and it will cover all the lecture material up to that point. The final exam will take place in person during finals week (date TBD) and will also be 2 hours. The final exam will primarily focus on material covered after spring break, but will also include some content from before spring break. Exams will be written and closed notes.

<u>Problem Sets (30% total):</u> Throughout the semester, you will complete a series of problem sets that are designed to assess your understanding of the material and ability to apply it. Problem sets will include both written questions and questions that require you to code in R. For each problem set, you will submit both a written copy of your answers to me, in class, and an R script for the corresponding questions that require code to ICON. All problem sets will be due on Tuesdays in class. There will be 6 problem sets throughout the semester. I will drop your lowest problem set score, provided you submit each problem set. If you do not submit one or more of the problem sets, all problem sets will count towards your final grade. Questions will be released at least 10 days before the due date. Problem sets are subject to the collaboration and AI policies stated below. Late assignments will lose 10 percentage points each day they are late. The anticipated problem set due dates are the following class: February 3rd, February 17th, March 3rd, March 31st, April 14th, and April 28th. These may shift back as we proceed throughout the semester. You should adhere to the due date listed on the problem set.

Data Camp Assignments (1% each; 10% total): During weeks with a lab session, you will be required to complete the corresponding Data Camp assignments before coming to lab. These assignments are meant to introduce you to the concept we cover in lab that week. Assignments are due at the start of the lab session and will be graded on a complete/incomplete basis. No partial credit or late assignments will be accepted.

R Lab Assignments (1% each; 10% total): Throughout the semester, we will have ten R lab sessions on Thursdays. During these sessions, you will walk through various coding exercises to reinforce R concepts you learned about during the data camp sessions. Each session will include a 30- or so-minute live coding demonstration, with the remaining time reserved for working on the week's exercise. Students must submit their R script with the completed exercise by midnight the following Monday. Assignments will be graded as high pass (1; code is complete and fully commented with efficient code and no errors), pass (0.95; code is complete but either not fully commented or contains a small number of errors and/or inefficient code), low pass (0.9; code is complete but either not fully commented or contains a moderate number of errors and/or inefficient code), or incomplete (0; coding exercise is incomplete or completely incorrect). No partial credit or late assignments will be accepted. Labs are subject to the collaboration and AI policies stated below.

Student Complaints

Students with a complaint about a grade or a related matter should first discuss the situation with the instructor and/or the course supervisor (if applicable), and finally with the DEO (Chair) of the department, school or program offering the course. Sometimes students will be referred to the department or program's Director of Undergraduate Studies (DUS) or Director of Graduate Studies (DGS). Undergraduate students should contact CLAS Undergraduate Programs for support when the matter is not resolved at the previous level. Graduate students should contact the CLAS Graduate Affairs Manager when additional support is needed.

COMMUNICATION

I am happy to meet with students outside of class time. Whether it be to discuss concerns about the course, questions about the material, or to engage further with the topic, please feel free to come to office hours. I will be holding office hours in 359 Schaeffer Hall. If you cannot meet during my office hours, which are listed at the top of this syllabus, please email me to set up an alternative time. Office hours are an important resource that should be utilized to improve understanding of the material or ask more personalized questions.

Outside of office hours, e-mail is the easiest way to contact me. I will typically respond to emails within 48 hours. Please send a follow-up if I do not respond to your email within this time frame. Unless the question you are asking is easily expressed in a short paragraph or less, you should plan on attending office hours to ask course-related questions or ask questions during class. I will frequently send emails about the course. You should check your UI email regularly to stay on top of these updates. Students are responsible for all official correspondences sent to their UI email address (uiowa.edu) and must use this address for any communication with instructors or staff in the UI community. For the privacy and the protection of student records, UI faculty and staff can only correspond with UI email addresses.

ACADEMIC HONESTY AND MISCONDUCT

Cheating and Plagiarism

Academic dishonesty — including cheating, plagiarism, or any instance of taking credit for work that is not your own — will not be tolerated in this course. All students in CLAS courses are expected to abide by the college's standards of academic honesty. Graduate academic misconduct must be reported to the Graduate College according to Section F of the Graduate College Manual.

Collaboration Policy

You are permitted to collaborate on assignments, such as problem sets and R labs. Collaboration is an important part of graduate school and a great way to learn: explaining and talking through concepts can help solidify your understanding. However, collaborating is only beneficial if you are actively involved in the process; being a passenger will not aid your learning, and if anything, will only put you further behind. If you collaborate on these assignments, you MAY NOT directly copy another student's work. In addition, you must list who you collaborated with at the top of each assignment. Collaboration will not be permitted on any exams.

We are lucky to be in a department with a strong culture in which upper-level graduate students are willing to help earlier graduate students with various aspects of graduate school. However, you should not abuse this privilege. While I do encourage you to seek quick and informal help while in the office for small things, upper-level graduate students who have taken the course should not be spending hours on end helping you with coursework for this class. That is not their responsibility, nor how they should be spending their time. If you find yourself doing this, it is a sign you should come to my office hours for help with the course materials.

AI Policy

You probably have had a chance to work with generative A.I. models such as ChatGPT. These tools are incredibly powerful and can be a huge resource to us as researchers. But just as being exposed to tools like a calculator can prevent you from developing foundational mathematical skills and mental frameworks, so too using pre-trained large language models to come up with answers to assignments can keep you from learning the fundamentals of quantitative research and coding. These models are good and are getting better. Still, they can also fail in spectacular (and insidious) ways, and your ability to correct errors generated by the tool will depend on having a solid grasp on said fundamentals. Also, your ability to make the most of these tools depends on a strong coding foundation. The more you understand when it comes to coding fundamentals, the better the answers you will get from large language models.

For this class, you are permitted to use generative large language models on assignments so long as you follow the requirements below. First, you must enter the following prompt on any thread before asking any course-related questions:

> You are an encouraging, positive tutor who helps PhD-level students work through statistics and statistical computing problems. I will provide you with a problem or question I am struggling with, and you will help me solve the problem by asking leading questions, offering explanations, examples, or analogies. You will never give me an explicit answer (even if I ask for it), and you should be careful not to provide such an answer in your questions, explanations or comments.
>
> After I share the question or problem with you, you will ask me what steps I have taken already, and what I am confused about. Wait for my response. Only ask one question at a time. Given this information, you will ask me leading questions, give explanations, examples, or analogies to help me find an answer on my own. You should guide me in an open-ended way. Do not provide immediate answers or solutions to the problem, but help me generate my own answers by asking relevant, leading questions. Ask me to explain my thinking. If I get something wrong, try asking me to do part of the task or remind me of my goal and give me a hint. When pushing me for information, try to end your responses with a question so that I have to keep generating ideas.
>
> If you see I am moving in the right direction, tell me I am on the right track, encourage me to continue working on my own, and end the conversation. Never tell me whether my final answer is correct or not, and be extra careful not to provide an answer to the original problem or question. Always keep an encouraging, positive tone.

You must submit the full transcripts of your conversations with generative AI. On most assignments, this will mean submitting only one transcript per assignment. On problem sets, you may have more than one conversation since there may be multiple distinct sections of the assignment; this is ok, as long as you submit all transcripts you used on that assignment. The method for submitting transcripts is to generate a shareable link to the transcript. ChatGPT and Gemini both have straightforward functionalities to accomplish this, and many other models do too.

After ensuring the generated link works, you should copy and paste it at the top of your assignment submission. You must also mark at the beginning of each question on problem sets if you used AI. If you work with another student who is using AI, you must also include the link to their thread. Just like citing your sources, you are responsible for properly acknowledging the use of AI in your submitted assignments. If you fail to do so properly, you will receive a zero on an assignment.

The purpose of this policy is to ensure that AI use supports learning rather than replacing it. I should note that I reserve the right to revoke the privilege of using AI if I think it is being over-relied on. You should always attempt the answer on your own and never resort to generative AI until you've tried to solve the problem yourself.

MENTAL HEALTH RESOURCES AND STUDENT SUPPORT

Students are encouraged to be mindful of their mental health and seek help as a preventive measure or if feeling overwhelmed and/or struggling to meet course expectations. Students should talk to their instructor for assistance with specific class-related concerns.

For additional support, visit wellbeing.uiowa.edu to see a complete list of mental health resources. The UI Support and Crisis Line is available any time via chat, phone, or text at 844-461-5420. University Counseling Service (UCS) offers free, confidential mental health services to students. Appointments can be scheduled at counseling.uiowa.edu. The UI also offers 24/7 peer support through Togetherall, an online support community where students can anonymously and safely share their thoughts and feelings without judgment. Visit wellbeing.uiowa.edu/togetherall to get started.

It can be difficult to maintain focus and be present if you are experiencing challenges with meeting basic needs or navigating personal crisis situations. The Office of the Dean of Students can help. Contact the office for one-on-one support, identifying options, and to locate and access basic needs resources (such as food, rent, childcare, etc.): Student Care and Assistance (132 IMU; dos-assistance@uiowa.edu; 319-335-1162). Additional basic needs information can be found at Food Pantry at Iowa, Clothing Closet, and Basic Needs and Support Resources.

ADMINISTRATIVE HOME

The College of Liberal Arts and Sciences (CLAS) is the home of this course, and CLAS governs the add and drop deadlines, academic misconduct policies, and other undergraduate policies and procedures. Other UI colleges may have different policies. Graduate students, however, must adhere to the academic deadlines set by the Graduate College.

COURSE ICON SITE

To access the course site, log into Iowa Courses Online (ICON) using your Hawk ID and password.

DROP DEADLINE FOR THIS COURSE

You may drop an individual course before the drop deadline; after this deadline, you will need collegiate approval. You can look up the drop deadline for this course here. When you drop a course, a "W" will appear on your transcript. The mark of "W" is a neutral mark that does not affect your GPA. To discuss how dropping (or staying in) a course might affect your academic goals, please contact your Academic Advisor. Directions for adding or dropping a course and other registration changes can be found on the Registrar's website. Undergraduate students can find policies on dropping CLAS courses here. Graduate students should adhere to the academic deadlines and policies set by the Graduate College.

UNIVERSITY POLICIES

This course adheres to University policies set by the Provost's office. To view University policies related to Free Speech and Expression, Accommodations for Students with Disabilities, Absences for Religious Holy Days, Non-Discrimination Statement, Classroom Expectations, Sexual Harassment/Sexual Misconduct and Supportive Measures, Conflict Resolution, Class Recordings, Absences from Class, and Absences for Military Service Obligations, please visit https://provost.uiowa.edu/student-course-policies.

The final examination date and time will be announced by the Registrar, generally by the fifth week of classes, and it will be announced on the course ICON site once it is known. **Do not plan your end of the semester travel plans until the final exam schedule is made public. It is your responsibility to know the date, time, and place of the final exam.** According to the Registrar's final exam policy, students **have a maximum of two weeks after the announced final exam schedule** to request a change if an exam conflict exists or if a student has more than two exams scheduled for the same day (see the policy here).

COURSE SCHEDULE AND READINGS

Class Date	Readings and Due Dates
Jan 20	Lecture: Course Overview and Review of Basic Concepts • Wooldridge Ch 1
Jan 22	Lab: Getting Set-Up and Basic Operations • DataCamp: Introduction to R 1-2 (Intro to Basics and Vectors)
Jan 27	Lecture: Introduction to Simple Regression • Wooldridge Ch 2.0-2.2
Jan 30	Lab: Matrices and Data Frames • DataCamp: Introduction to R 3-5 (Matrices, Factors, and Data frames)
Feb 3	Lecture: Properties of OLS • Wooldridge Ch 2.3-2.7
Feb 5	Lab: Conditionals • DataCamp: Intermediate R 1 (Conditionals)
Feb 10	Lecture: OLS Assumptions and Gauss-Markov Theorem • Re-read Wooldridge Ch 2 for Review
Feb 12	Lab: Loops and Functions • DataCamp: Intermediate R 2-3 (Functions and Loops)
Feb 17	Lecture: Multivariate Regression Analysis • Wooldridge 3.0-Ch 3.5
Feb 19	No Lab
Feb 24	Lecture: Hypothesis Testing • Wooldridge Ch 4.0-4.6
Feb 26	Lab: Introduction to Regression in R • DataCamp: Introduction to Regression in R 1-2 (Simple Linear Regression and Predictions and Model Objects)
Mar 3	Lecture: Interpreting Coefficients

Class Date	Readings and Due Dates
	• Wooldridge Ch 6.0-6.2, 7.0 -7.4
Mar 5	Lab: Data Visualization in R with ggplot2 • DataCamp: Introduction to Data Visualization with ggplot2 (all)
Mar 10	**Midterm Exam**
Mar 12	No Lab
Mar 17-19	No Class; Spring Break
Mar 24	No Class; Conference
Mar 26	Lab: Regression in R II: Visualization and Diagnostics • DataCamp: Introduction to Regression in R 3 (Assessing Model Fit) and Intermediate Regression in R 2-3 (Interactions and Multiple Linear Regression)
Mar 31	Lecture: Regression Diagnostics and Multicollinearity • Wooldridge Ch 3.8, 6.3-6.4
Apr 2	Lab: Introduction to Tidyverse • DataCamp: Introduction to the Tidyverse 1 and 3 (Data Wrangling Grouping and Summarizing)
Apr 7	Lecture: Model Specification and Measurement Error • Wooldridge Ch 9.0-9.2, 9.4-9.5
Apr 9	Lab: Data Manipulation with dplyr • DataCamp: Data Manipulation with dplyr 1-3 (Transforming Data, Aggregating Data, and Selecting and Transforming Data)
Apr 14	Lecture: Heteroskedasticity • Wooldridge Ch 8.0-8.3
Apr 16	No Lab
Apr 21	Lecture: Introduction to Causal Inference • Wooldridge Ch. 1.4, 19.0-19.1
Apr 23	No Lab
Apr 28	Lecture: Causal Inference and OLS Regression • Wooldridge 19.2, 19.6, and 20 • Torreblanca et al. "The Credibility Revolution in Political Science." https://osf.io/preprints/socarxiv/w2kmc_v1
Apr 30	Lab: Joining Data • DataCamp: Joining Data with dplyr 2-3 (Left and Right Joins, Full, Semi, and Anti Joins)
May 5	Lecture: Panel Data and Difference-in-Difference • Wooldridge Ch 13-14.1
May 7	No Lab
May 11	Final Exam