Public Policy 571: Applied Econometrics  
Fall 2023 Syllabus

Instructor: Jonathan Hanson (he/him)  
jkhanson@umich.edu  
4223 Weill Hall, 734-615-1496  
Office Hours: Mon. 3:00–4:30, Fri. 11:00–12:00, or by appointment

The objective of this course is to help gain students proficiency in advanced techniques that are employed in the service of causal inference, including selection models, regression discontinuity, difference-in-differences, matching methods, and interrupted time series. Students will understand how to design, execute, and interpret the results from estimation techniques beyond the standard OLS model. Students will learn to work with data structures in which the observations are not independently and identically distributed: time-series data, panel data, time-series-cross-sectional data, multi-level data, etc.

Prerequisite: it is assumed that students in this course will have previously taken Public Policy 529 and Public Policy 639, or equivalent coursework. Students are advised to consult with the instructor if they received a grade lower than A- in either of these courses.

Class Meeting Schedule

Unless otherwise noted, lectures are Mondays and Wednesdays from 10:00–11:20 pm in 1230 Weill Hall.

Textbooks

There is no single textbook for this course, and readings will be available on the Canvas website. You can log into Canvas at http://canvas.umich.edu. The following books and articles, however, will be good references for course material. The Cunningham and Huntington-Klein texts are available in an online format.

Assignments and Grading

Your grade for this course will be determined by the following:

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem sets</td>
<td>50%</td>
</tr>
<tr>
<td>Midterm exam</td>
<td>25%</td>
</tr>
<tr>
<td>Final exam</td>
<td>25%</td>
</tr>
</tbody>
</table>

There will be six major problem sets in this course. Additionally, there will be two exams – one midterm and one final – covering the first-half and second halves of the course respectively. Each exam will have two components: a take-home part and an in-class part.

You are encouraged to help each other figure out the answers to the problem sets, but it is expected that you write up your answers independently. The take-home exams are exams. For these exams, you are not permitted to communicate with any person other than the course instructor. Receiving help from any other person, or providing help to any other person, is a violation of academic integrity and will result in significant consequences. Use of generative AI tools to answer exam questions is also an academic integrity violation. You may, however, consult any textbook or static internet-resources.

Assignment and Exam Calendar

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem Set 1</td>
<td>September 18</td>
</tr>
<tr>
<td>Problem Set 2</td>
<td>September 27</td>
</tr>
<tr>
<td>Problem Set 3</td>
<td>October 9</td>
</tr>
<tr>
<td>Take-home Midterm Exam</td>
<td>October 23</td>
</tr>
<tr>
<td>Problem Set 4</td>
<td>November 6</td>
</tr>
<tr>
<td>Problem Set 5</td>
<td>November 20</td>
</tr>
<tr>
<td>Problem Set 6</td>
<td>December 8</td>
</tr>
<tr>
<td>Final Exam (in-class portion)</td>
<td>December 13</td>
</tr>
<tr>
<td>Final Exam (take-home portion)</td>
<td>December 18</td>
</tr>
</tbody>
</table>

The final course letter grade reflects the Ford School’s guidelines. An A is awarded for work that is Excellent, an A- for work that is Very Good, a B+ for work that is Good, a B for work that is Acceptable, and a B- for work that is below expectations for graduate work. You should know I do not have a predetermined formula to convert numeric point totals into these categories. It would be a mistake, for instance, to assume that a grade of 75% on an exam translates into a C, since exams vary in their difficulty.

Software

Students can use either R or Stata statistical software to complete homework assignments. These applications are available on computers in the Ford School computer lab, as well as the larger com-
puter labs on campus. Additionally, students can remotely log in to the university’s Virtual Sites (see information at https://its.umich.edu/computing/computers-software/virtual-sites) to access Stata and other university site-licensed software when not on campus.

R is an open-source program that is freely downloadable from https://cran.r-project.org. Students who use R are strongly encouraged to download the free RStudio Desktop companion application (https://rstudio.com/products/rstudio/download/) to serve as their interface with R. Time-limited licenses for Stata are available for purchase at student pricing from https://www.stata.com/order/new/edu/profplus/student-pricing/. The Stata/BE version is sufficient for this course.

There are several resources for learning Stata available on Canvas, including a handbook that I compiled for Public Policy 567. On Canvas, you will also find an R handbook developed specifically for this course. When possible, the handbook replicates any examples from lecture, and it provides some coverage of R-specific tools that do not appear in Stata.

Additionally, Alton Worthington, a lecturer here at the Ford School who teaches several software-related courses, holds general office hours every Wednesday from 10:00-4:00. He is happy to help anyone with coding-related questions. You can sign up for these office hours at http://bit.ly/ABHW-00. Meetings can be in his office (3216 Weill Hall) or via Zoom.

**Academic Integrity**

The Ford School academic community, like all communities, functions best when its members treat one another with honesty, fairness, respect, and trust. We hold all members of our community to high standards of scholarship and integrity. To accomplish its mission of providing an optimal educational environment and developing leaders of society, the Ford School promotes the assumption of personal responsibility and integrity and prohibits all forms of academic dishonesty, plagiarism and misconduct. Academic dishonesty may be understood as any action or attempted action that may result in creating an unfair academic advantage for oneself or an unfair academic advantage or disadvantage for any other member or members of the academic community. Plagiarism involves representing the words, ideas, or work of others as one’s own in writing or presentations, and failing to give full and proper credit to the original source. Conduct, without regard to motive, that violates the academic integrity and ethical standards will result in serious consequences and disciplinary action. The Ford School’s policy of academic integrity can be found in the MPP/MPA, BA, and PhD Program handbooks. Additional information regarding academic dishonesty, plagiarism and misconduct and their consequences is available here.

**Inclusivity**

Members of the Ford School community represent a rich variety of backgrounds and perspectives. We are committed to providing an atmosphere for learning that respects diversity. While working together to build this community we ask all members to:

- share their unique experiences, values and beliefs
- be open to the views of others
- honor the uniqueness of their colleagues
• appreciate the opportunity that we have to learn from each other in this community
• value one another’s opinions and communicate in a respectful manner
• keep confidential discussions that the community has of a personal (or professional) nature
• use this opportunity together to discuss ways in which we can create an inclusive environment in Ford classes and across the UM community

Accommodations for Students with Disabilities

The University of Michigan recognizes disability as an integral part of diversity and is committed to creating an inclusive and equitable educational environment for students with disabilities. Students who are experiencing a disability-related barrier should contact Services for Students with Disabilities (SSD); 734-763-3000 or ssdoffice@umich.edu). For students who are connected with SSD, accommodation requests can be made in Accommodate. If you have any questions or concerns please contact your SSD Coordinator or visit SSD’s Current Student webpage. SSD considers aspects of the course design, course learning objects and the individual academic and course barriers experienced by the student. Further conversation with SSD, instructors, and the student may be warranted to ensure an accessible course experience.

Student Mental Health and Wellbeing

The University of Michigan is committed to advancing the mental health and wellbeing of its students. We acknowledge that a variety of issues, both those relating to the pandemic and other issues such as strained relationships, increased anxiety, alcohol/drug problems, and depression, can directly impact students’ academic performance and overall wellbeing. If you or someone you know is feeling overwhelmed, depressed, and/or in need of support, services are available.

You may access the Ford School’s embedded counselor Paige Ziegler (pziegler@umich.edu) and/or counselors and urgent services at Counseling and Psychological Services (CAPS) and/or University Health Service (UHS). Students may also use the Crisis Text Line (text ‘4UMICH’ to 741741) to be connected to a trained crisis volunteer. You can find additional resources both on and off campus through the University Health Service and through CAPS.

Ford School Public Health Protection Policy

In order to participate in any in-person aspects of this course—including meeting with other students to study or work on a team project—you must follow all the public health safety measures and policies put in place by the State of Michigan, Washtenaw County, the University of Michigan, and the Ford School. Up to date information on U-M policies can be found here. It is expected that you will protect and enhance the health of everyone in the Ford School community by staying home and following self-isolation guidelines if you are experiencing any symptoms of COVID-19.

Use of Technology

Students should follow instructions from their instructor as to acceptable use of technology in the classroom, including laptops, in each course. All course materials (including slides, assignments,
handouts, pre-recorded lectures or recordings of class) are to be considered confidential material and are not to be shared in full or part with anyone outside of the course participants. Likewise, your own personal recording (audio or video) of your classes or office hour sessions is allowed only with the express written permission of your instructor. If you wish to post course materials or photographs/videos of classmates or your instructor to third-party sites (e.g. social media), you must first have informed consent. Without explicit permission from the instructor and in some cases your classmates, the public distribution or posting of any photos, audio/video recordings or pre-recordings from class, discussion section or office hours, even if you have permission to record, is not allowed and could be considered academic misconduct.

Advances in generative artificial intelligence (AI) tools, such as ChatGPT, create a wide range of challenges and opportunities for higher education. While these tools can be extremely helpful in many respects, they do not change a fundamental principle of academic work: it is wrong to claim someone else’s ideas or work as your own. For a student, the downsides and upsides of working with an AI tool can be much like working with a study group. Working with others can help you overcome roadblocks and help you learn things that you may have missed, but over-reliance on others to obtain answers likely will inhibit your learning. For data analysis, we have many tools at our disposal. For example, ultimately we will use statistical software to do almost all the difficult work for us, but it is critical that we understand what the software is doing and that we know what to tell the software to do.

Here are a few guidelines for using generative AI tools. If you have any questions about whether a particular use is permissible, you should ask your instructor.

- It is okay to use these tools as a resource to look up information about theoretical concepts. Students should be aware that the information they receive may not be accurate or may not match what they are taught in lectures or their course textbook.
- It is okay to use these tools as a resource to help with learning the coding syntax of R or Stata. These tools can be very helpful with sample code.
- It is not permissible to use an AI tool to answer assignment questions for you. If you do use AI for assistance on how to do a particular kind of question, you should cite the query that you supplied. Please note that these tools make mistakes. You are responsible for any errors in the answer you submit.
- On the take-home midterm or final exams, AI tools can only be used in manner equivalent to that of a textbook or generic internet resource. Querying an AI tool about how to answer a particular question on the exam is a serious academic integrity violation equivalent to communicating with another person about that question.

Please review additional information and policies regarding academic expectations and resources at the Ford School of Public Policy at: https://intranet.fordschool.umich.edu/academic-expectations.
Advanced Estimation Methods for Cross-Sectional Data

August 28: Principles of Maximum Likelihood Estimation

In the classical linear regression model and its variants, we find the parameters of interest (i.e. the regression coefficients) through analytic formulas. With Maximum Likelihood Estimation (MLE), these parameters are often found by a search algorithm which tours the parameter space and finds the values of the parameters that are “most likely” given the data. The power of MLE its flexibility. Many different functional forms are possible, facilitating analysis of many kinds of data for which linear regression is not suitable.


August 30 & September 6: Dichotomous Dependent Variables

When our dependent variables are dichotomous, OLS might be problematic. For one thing, OLS can in some cases produce predicted values of the dependent variable that fall outside the range the 0-1 range. Other problems arise with non-normality of disturbances and non-linear effects of independent variables. With Maximum Likelihood Estimation, we can use functional forms that are designed for these data: probit and logit models.


September 11: Ordinal Categories as Dependent Variables

When the dependent variable consists of ordinal categories, the list of possible problems grows. We can use ordered probit and ordered logit models for this kind data. A key objective is to learn how to interpret and convey the results.

September 13: Nominal Categories as Dependent Variables

With unordered categories, estimation becomes trickier since we cannot use a latent dimension as an underlying framework. We can use multinomial logit models in these circumstances.


September 18 & 20: Count Models

Count models are for cases in which the dependent variable is a count of the number of times something occurs.


Causal Inference for Cross-Sectional Data

September 25 & 27: Sample Selection Models and Related Models

The purpose of selection models is to address situations in which the cases that make it into the sample are different in important, unmeasured ways from those who do not, and these unmeasured factors are relevant for predicting the dependent variable. In other situations, assignment into the treatment condition is non-random and correlated with relevant, unmeasured factors.

October 4 & 5: Multivariate Matching Methods

With observational data, we face inherent challenges in estimating the average treatment effect (ATE) of some policy intervention because the treatment is not randomly-assigned. The treatment group and control group may differ from each other, in the aggregate, on both observable and unobservable characteristics in ways other than the treatment, and we cannot adequately control for these other factors. Matching methods are intended to bring balance to the treatment groups on these other characteristics to reduce bias in estimating the ATE. In this section of the course we will learn various matching methods and develop awareness of their limitations.


October 9 & 11: Regression Discontinuity Models

Regression discontinuity models are useful for scenarios in which assignment into the treatment group is based upon a cutoff score on some observable characteristic or a particular date of implementation. In short, the difference between the predicted outcomes on either side of the cutoff point becomes a way to estimate the size of the treatment effect. This section of the course explores various uses of these models.

October 16: Fall Study Break

October 18: Instrumental Variables Analysis

Oftentimes, there is strong reason to believe that the relationship between our dependent variable, and one or more of our independent variables, is endogenous. In such cases, our estimate of the coefficient on the independent variable will be biased. If we can find a third variable, however, which only has an effect on the dependent variable through the endogenous independent variable, we can identify the correct relationship.

Estimation Methods for Time-Series and Panel Data

October 23 & 25: Regression with Time-Series Data

The standard assumption is that our data are independent and identically distributed. When we have time-series data, such as the results of monthly polls on presidential approval or some other case of repeated observations of the same object, this assumption is violated. The stochastic component of one observation may be correlated with the one preceding it and the one that follows. We spend about two sessions on this topic.
October 30: Multi-Level Models

When our data consist of individual cases that are embedded within higher-level units – such as school children inside classrooms, which are inside schools – we need methods that can estimate both individual and unit-level effects. In this section of the course, we explore the use of multi-level models for this purpose.


November 1, 6 & 8: Regression with Panel Data

We have a panel when our data contain repeated observations of a sample of objects, such as a set of individuals who are surveyed periodically or a time series of cross-national data. In this scenario, we need to think about the non-independence of our observations both across time and space.

Causal Inference for Time-Series and Panel Data

November 13, 15 & 20: Difference-in-Difference Models

Difference-in-difference models are useful when: 1) we have observations of each of the members of the sample for at least two periods in time; and, 2) one group within the sample experienced the treatment between these two periods, while the other did not. We can thus observe whether group-level differences in the outcome of interest changed from one period to the next, facilitating an estimate of the treatment effect.


November 27 & November 29: Interrupted Time Series

With cross-sectional data, we used regression discontinuity models to make a causal inference around a cutoff point on an observable characteristic. With time-series data, that cutoff point can be a point in time, such as the date of a policy implementation. Comparison cases facilitate the identification of shifts in levels and trends of the outcome variable.


December 4 & 6: Synthetic Control Method

This method builds upon matching methods by creating a “synthetic” comparison case out of a pool of possible cases. Ideally, we would be able to compare a case that has undergone some treatment to the counterfactual scenario in which it had not undergone the treatment. Since this counterfactual scenario does not exist, we construct one using a weighted average of similar cases that did not undergo the treatment. We then can compare the observed outcomes against those predicted by under the counterfactual.


In-class portion of Final Exam Wednesday, December 13, 10:30-12:30

Take-home portion of Final Exam due by Monday, December 18