Causal Inference and Research Design

University of Texas at Austin, 395M

Monday and Wednesday,

BRB 1.118 1:00pm-3:30pm

Professor and TA information

Dr. Scott Cunningham
Preferred email address: scunning@gmail.com
https://www.scunning.com
Office Location: TBD
Office Hours: Monday and Wednesday 10:30pm - 12:00pm or by Zoom appointment
Text: (254) 537-2239
Github repo: `https://github.com/scunning1975/UT-Causal-Inference`
Platform: `https://www.twitch.tv/causalinf_did`
Newsletter: `https://causalinf.substack.com`
TA: Nathan Hattersley (nhattersley@utexas.edu)
Calendy: `https://calendly.com/nhattersley`

This syllabus may change, but I am making an effort to make it fixed.

Course Description: The modern field of causal inference is an applied statistics sub field within econometrics built on a theory of counterfactuals and the potential outcomes framework (Cunningham, 2021; Imbens and Rubin, 2015). My course is primarily based on the design tradition, though I will also sometimes discuss the model tradition too (Pearl, 2009). Our focus this semester is to train you in the art and science of this field by learning about empirical microeconomics, the most popular research designs and econometric estimators used within design based causal inference. It is also to further you along in your knowledge of how to use software and programming languages to conduct this empirical work.

Prerequisite(s): Econometrics or equivalent.

Course Objectives:

- to develop a solid understanding of design based causal inference.
- to develop a solid understanding of empirical workflow
- to develop a solid understanding of statistical language (like R, python or julia) or statistical software (like Stata)
- to gain experience applying all three to empirical microeconomics research

How I will try to help you reach these goals:

1. **Coding together in class**. Every now and then throughout the semester, we will take breaks and work on the Lott and Mustard (1997) project, as well as your research project. You are required on those days to participate in our coding together collaborations and **cannot** work on other work.

2. **Empirical workflow**. You will be early on doing some simple analysis related to data collection, data cleaning, automation and production of tables and figures, as well as writing a simple paper discussing what you found. Empirical workflow will be part of your grade.

3. **Replications**. We will apply these research designs throughout the semester. The difference-in-differences replication is semester long and will go through stages of development, but the others will be based on what we are covering around that time.

4. **Original empirical project**. You will collect your own data by the week before spring break (last week of February) and show me simultaneously an idea, a design/strategy, and preliminary analysis. You cannot proceed unless all three occur.

Required and Supplemental Textbooks

There is only one required textbook, but you do not have to purchase it as it is available online free. There is also an excellent supplemental book I am assigning

- **Required**: Cunningham (2021), Causal Inference: the Mixtape (Yale University Press). This can be accessed via my website for free at `https://mixtape.scunning.com`.
- **Supplemental**: Huntington-Klein (2022), The Effect: An Introduction to Research Design and Causality (CRC Press). This can be accessed for free at `https://theeffectbook.net`.

Classwork and your final grade:

- **Coding assignments (25%)**: Throughout the semester, there will be several coding assignments. You are expected to work on these together but you must turn in your assignments separately. These are pass/fail. The coding assignment includes a "hidden curriculum", or what I will sometimes call "empirical workflow" assignment. My TA, Nathan Hattersley, will be teaching a three week (starting Friday January 21st) mini-course on R, version control, and basic data wrangling. This will be the foundation for you to learn R, but also github and basic data manipulation. An assignment will be due mimicking what you do as a group in that mini-course. Attendance is mandatory. The Github repo for Nathan's sequence is `https://github.com/nateybear/causal-inference-2022/`.

- **Extending Lott and Mustard (1997) (25%)**: A main part of the semester is an extension to a controversial paper on concealed carry gun laws from the 1990s by Lott and Mustard (1997). You are expected to be thoroughly familiar with the original paper, the followup criticism by Donohue, Aneja and Zhang (2011), the laws more generally, the original methods the authors used, and the datasets too. The datasets are a state-level version and a county-level version. We will be analyzing both at all times so that we can better understand the effect, if anything, that these laws had on crime outcomes. See `https://works.bepress.com/john_donohue/89/` and scroll to the bottom for the data.

  The main way that I am trying to get us to make progress on this assignment is, in addition to your private time on the project, through our collaborative group meetings I sometimes call "coding together". In addition to extending the analysis using more contemporary difference-in-differences estimators, we are also trying to better understand the problems that can occur during the data preparation stage and solve them ahead of time. The county-level data in particular has been criticized and we must clean and prepare it. If we do not follow principled workflow principles, it can lead to large variation in data construction across researchers, and a secondary goal of this assignment is to understand why that happens and how to minimize it. Thus we will code together regularly to discuss everyone's progress as we try to reconcile differences in our analysis as well as to extend the paper using the Callaway and Sant'Anna (2020) estimator and the stacking method. We will meet regularly throughout the semester to discuss what we are finding, solve problems, "code together", and reconcile results. Our goal by early April is to have completed our analysis and produced a final review of both the county and state level analysis using TWFE, Callaway and Sant'Anna (2020) and stacking. But our goal is to learn

together, collaborate, and deepen our understanding of things that can cause analysis to go wrong, not by coding errors during estimation, but also data preparation before analysis. You are encouraged to work together even though your final output will be a paper written in the style of an empirical microeconomics paper discussing the original paper, the theory, the dataset, the methodological issues with the original paper, your two extensions (in addition to a replication), results and interpretation and conclusion.

- **Original research project (50%)**: The primary input in your final grade is an original empirical microeconomics research paper. This project will require that you collect data, clean and prepare it using the principles discussed in Nathan's mini-course, analyze the data using a research design that can answer your question and an appropriate econometric estimator. It will require writing a formal paper in the style and rhetoric of an empirical microeconomics paper. You will be judged on success which includes collecting the data, successfully cleaning it following the empirical workflow set up this semester, analysis using research design and econometric modeling, and a good paper following the same structure I just outlined.

Github repository I will be distributing materials through Github this semester, not Canvas. This is because Github (and git) is the dominant way in which empirical work is done today, but also it is the way in which new estimators are distributed as increasingly statisticians and economists only write in R. Stata distributes its new packages via ssc, but usually R authors distribute new packages via Github. Our Github repo is `https://github.com/scunning1975/University-Causal-Inference`.

Credit and grade distribution Students will be evaluated based coding assignments, extension project, and original research project.

Grade Distribution:

Coding assignments	25%
Extension of Lott and Mustard (1997) paper	25%
Original research project	50%

Final grade scales:

>= 90-100	A	70 - 89.99	B
50 - 69.99	C+	<= 49.99	F

COVID related policies

As each of you know, the United States is now two years into a global pandemic. As of this writing, over 850,000 Americans have died from COVID and 5.55 million worldwide. The epidemic has moved into its second variant called Omicron which is more contagious than the past ones. This disease has been a historic once-in-a-lifetime event affecting public health, mortality, the functioning of our economy, as well as school enrollment and pedagogy. There are two things we must discuss, therefore, related to the pandemic: the format of the class which can change over the semester depending on local factors and mandatory masking. I will discuss each of these now

Class Format Conversion

Our class will meet in person this semester. But I will be recording zoom lectures too and distribute them after class asap. But because of Omicron, it's possible that we have to switch to an online format. If that happens, then we will meet at the scheduled time over Zoom. Regardless, though, the zoom lectures cannot be shared by UT policy. Class recordings are reserved only for students in this class for educational purposes and are protected under FERPA. The recordings should not be shared outside the class in any form. Violation of this restriction by a student could lead to Student Misconduct proceedings.

Mandatory Masking

University of Texas is committed to providing a safe educational environment for all students, faculty, and staff. In order to create the healthiest environment possible, students and instructors must properly wear face coverings over their mouths and noses while in University of Texas buildings and while in the classroom. In addition, social distancing should be maintained within all classroom environments.

Any student who does not bring a face covering or refuses to wear one will be dismissed from class and be considered absent which for our class means you will not receive class participation for that day. If a student who is not wearing a face covering refuses to leave class, the instructor may dismiss in-person class for that day. Such a student will be subject to disciplinary action through the Student Code of Conduct.

Lauren's Promise Lauren McCluskey, a 21-year-old honors student athlete at the University of Utah, was murdered on Oct. 22, 2018, by a stalker. We must all take actions to ensure this never happens again. Lauren's mother has asked professors to consider adding to the syllabus an assurance of assistance to students in danger called Lauren's Promise. **Lauren's Promise** is my promise to "listen and believe you if someone is threatening you." Any form of sexual harassment or violence will not be excused or tolerated at University of Texas.

In cases of sexual harassment or violence, University of Texas will:

- Respond promptly and effectively to sexual assault, relationship violence, and stalking,
- Provide interim measures as necessary
- Provide confidential and non-confidential support resources,
- Conduct a thorough, reliable, and impartial investigation through its Title IX office
- Provide remedies as necessary.

If you are experiencing sexual assault, relationship violence, or stalking, you can take the following actions:

- If you are in immediate danger, call 911.
- Call University of Texas Police Department at (512) 471-4441
- Report it to me, and I will connect you to resources. Please call or text me at any hour using my cell phone which is (254) 537-2239. I am pretty good at immediately responding. You can also DM me on Twitter (even anonymously if you prefer) at @causalinf.
- Seek confidential sources of support and help:
- Counseling and Consultation Service Available Monday to Friday 8am to 5pm through UT's Counseling Center, `https://cmhc.utexas.edu`, (254) 710-2467
- Health services, Monday to Friday, 8am - 5pm, `https://healthyhorns.utexas.edu/appointments.html`, or you can call the crisis line at (512) 471-2255, or schedule an appointment between the hours 8:00am and 5:00pm Monday to Friday by calling (512) 471-3515
- Report the offense to the Title IX Office using the following URL (click on "Make a Title IX Report" at the top of the page at `https://titleix.utexas.edu/file-a-report`), or email them using the following email address titleix@austin.utexas.edu, or call them at (512) 471-0419

Title IX Office

Beginning January 1, 2020, Texas Senate Bill 212 requires all employees of Texas universities, including faculty, report any information to theTitle IX Office regarding sexual harassment, sexual assault, dating violence and stalking that is disclosed to them. Texas law requires that all employees who witness or receive any information of this type (including, but not limited to, writing assignments, class discussions, or one-on-one conversations) must be reported. If you would like to speak with someone who can provide support or remedies without making an official report to the university, please email advocate@austin.utexas.edu. For more information about reporting options and resources, visit `http://www.titleix.utexas.edu/`, contact the Title IX Office via email at titleix@austin.utexas.edu, or call (512) 471-0419. Although graduate teaching and research assistants are not subject to Texas Senate Bill 212, they are still mandatory reporters under Federal Title IX laws and are required to report a wide range of behaviors we refer to as sexual misconduct, including the types of sexual misconduct covered under Texas Senate Bill 212.The Title IX office has developed supportive ways to respond to a survivor and compiled campus resources to support survivors.

Makeup exam policy My makeup exam policy is that if you miss an exam for any reason, then the final exam will count for both the missed exam and the final itself.

Accessible, Inclusive, and Compliant Statement If you are a student with a disability, or think you may have a disability, and need accommodations please contact Services for Students with Disabilities (SSD). You may refer to SSD's website for contact and more information: `http://diversity.utexas.edu/disability/`. If you are already registered with SSD, please deliver your Accommodation Letter to me as early as possible in the semester so we can discuss your approved accommodations.

The university is committed to creating an accessible and inclusive learning environment consistent with university policy and federal and state law. Please let me know if you experience any barriers to learning so I can work with you to ensure you have equal opportunity to participate fully in this course. If you are a student with a disability, or think you may have a disability, and need accommodations please contact Services for Students with Disabilities (SSD). Please refer to SSD's website for contact and more information: http://diversity.utexas.edu/disability/. If you are already registered with SSD, please deliver your Accommodation Letter to me as early as possible in the semester so we can discuss your approved accommodations and needs in this course.

Academic Honesty

Students who violate University rules on academic dishonesty are subject to disciplinary penalties, including the possibility of failure in the course and / or dismissal from the University. Since such dishonesty harms the individual, all students, and the integrity of the University, policies on academic dishonesty will be strictly enforced. For further information, please visit the Student Conduct and Academic Integrity website at: `http://deanofstudents.utexas.edu/conduct`. You can find a link to the Standards of Conduct page at `https://deanofstudents.utexas.edu/conduct/standardsofconduct.php`.

Tentative Schedule and Topics

The following is a rough schedule of the topics we will cover in class. Usually, the Oyer chapters will be due every Thursday until we run out of chapters. This is a tentative schedule, but I will attempt to stay on schedule. The first day of class is January 19th and runs through May 5th 2021. We will start the first day immediately with lectures. I will be using a deck of slides in pdf form which I will be posting to Canvas. You are free to download them obviously and use them to help you understand the lectures.

Tentative Course Outline:

The weekly coverage might change as it depends on the progress of the class. However, you must keep up with the reading assignments. Remember each Friday your Substack must cover no fewer than 2 of the readings.

Week	Content
Week 1	• Review syllabus, history of design based causal inference, hidden curriculum, empirical workflow • **Due Friday night by midnight**: Personal essay
Week 2	• Finish Hidden curriculum, empirical workflow • Mixtape Chapter 4: Potential outcomes
Week 3	• Mixtape Chapter 4: Potential outcomes • Effect Chapter 10-11 • **Due Friday night by midnight**: Perfect doctor exercise
Week 4	• Mixtape Chapter 8: Panel (Two-way fixed effects or TWFE) • Effect Chapter 16 • **Due Friday night by midnight**: Empirical workflow assignment
Week 5	• Mixtape Chapter 9: Difference-in-differences – regression, event studies and Abadie and Gardeazabal (2003), Sant'Anna and Zhao (2020) • Effect Chapter 18 • Coding Together on Thursday – TWFE, cleaning, focus on sample sizes for Lott and Mustard (1997)
Week 6	• Mixtape Chapter 9: Difference-in-differences and differential timing Callaway and Sant'Anna (2020), Goodman-Bacon (2021), Sun and Abraham (2020) • Effect Chapter 18
Week 7	• Mixtape Chapter 9: Difference-in-differences and differential timing – Cengiz et al. (2019) • Effect Chapter 18 • Coding Together on Thursday – implementing stacking (anticipate a lot of problem solving), estimating Callaway and Sant'Anna (2020)
Week 8	• Mixtape Chapter 6: Regression discontinuity • Effect Chapter 20 • Coding Together on Thursday – continue figuring out stacking and using Callaway and Sant'Anna (2020)
Week 9	• **SPRING BREAK – March 14 to 19**

Week	Content
Week 10	• Mixtape Chapter 6: Regression discontinuity • Effect Chapter 20 • Coding Together on Thursday: Work on DiD and/or Hansen (2015) replication
Week 11	• Mixtape Chapter 7: Finish RDD; Instrumental variables • Effect Chapter 19
Week 12	• Mixtape Chapter 7: Instrumental variables • Effect Chapter 19 • **Due Friday night by midnight**: Extension of Lott and Mustard (1997) (final draft)
Week 13	• Mixtape Chapter 7: Instrumental variables • **Due**: Hansen (2015) replication
Week 14	• Mixtape Chapter 10: Synthetic control
Week 15	• Mixtape Chapter 5: Matching • Effect Chapter 14
Week 16	• DAGs if we have time • **Due on Friday by midnight**: Synthetic control assignment • **Original research project due on Friday, May 6th**

Articles and Books on the Syllabus

Abadie, Alberto and Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." American Economic Review 93(1):113–132.

Callaway, Brant and Pedro H. C. Sant'Anna. 2020. "Difference-in-differences with Multiple Time Periods." Journal of Econometrics .

Cengiz, Doruk, Arindrajit Dube, Attila Lindner and Ben Zipperer. 2019. "The Effect of Minimum Wages on Low-Wage Jobs." Quarterly Journal of Economics pp. 1405–1454.

Cunningham, Scott. 2021. Causal Inference: The Mixtape. 1st ed. Yale University Press.

Donohue, John J., Abhay Aneja and Alexandria Zhang. 2011. "The Impact of Right to Carry Laws and the NRC Report: Lessons for the Empirical Evaluation of Law and Policy." American Law and Economics Review 13(2):565–632.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." Journal of Econometrics .

Hansen, Ben. 2015. "Punishment and Deterrence: Evidence from Drunk Driving." American Economic Review 105(4):1581–1617.

Huntington-Klein, Nick. 2022. The Effect: An Introduction to Research Design and Causality. 1st ed. CRC Press.

Imbens, Guide W. and Donald B. Rubin. 2015. Causal Inference for Statistics, Social and Biomedical Sciences: An Introduction. 1st ed. Cambridge University Press.

Lott, John R. and David B. Mustard. 1997. "Crime, Deterrence and the Right-to-Carry Concealed Handguns." Journal of Legal Studies 26:1–68.

Pearl, Judea. 2009. Causality. 2nd ed. Cambridge University Press.

Sant'Anna, Pedro and Jun Zhao. 2020. "Doubly Robust Difference-in-Differences Estimators." Journal of Econometrics 219:101–122.

Sun, Liyang and Sarah Abraham. 2020. "Estimating Dynamic Treatment Effects in Event Studies with Heterogenous Treatment Effects." Journal of Econometrics Forthcoming.