**Name:** Scott Cunningham (Professor of Economics, Baylor University)

**Website:** https://www.scunning.com

**Email:** [scunning@gmail.com](mailto:scunning@gmail.com)

**Substack:** <https://causalinf.substack.com>

**Textbook**: <https://mixtape.scunning.com>

**Twitter**: @causalinf

**Github repo:**

<https://github.com/scunning1975/CodeChella_DiD>

**Slides**

./Slides/codechella.pdf

./Slides/workshop.pdf

**Assignments**

./Assignments

**Website with DiD programs in R and Stata**

<https://asjadnaqvi.github.io/DiD/>

**About the instructor:**

Scott Cunningham is a professor of economics at Baylor University in Waco Texas. He has published in top economics outlets such as *The Review of Economic Studies, Journal of Urban Economics, Journal of Human Resources, Journal of Public Economics, Journal of Development Economics* and more. He is the author of Causal Inference: the Mixtape published by Yale University Press in 2021 and co-editor of The Handbook on the Economics of Prostitution (with Manisha Shah) published by Oxford University Press in 2016. His research focus covers a range of applied topics in health and labor, including sex work, abortion, drug policy and mental healthcare. He has taught dozens of in-person and online workshops on causal inference and difference-in-differences to universities and firms across the world including Facebook, HP, University of Oxford, London School of Economics, University of Pennsylvania and many more.

**Description of class:**

Causal inference is a specialization within economics and statistics that grew out of the labor economics tradition to evaluate the causal effects of programs. While physical randomization was widely known to yield unbiased estimates of causal effects, it was not often used in economics. A wave of new labor economists starting in the late 1970s and mid 1980s changed that as they pushed for focus on exploiting quasi random assignment in natural experiments or through imposed modeling assumptions on counterfactuals. This work in conjunction with pioneering work in econometrics led to the sharpening of such research designs as instrumental variables and difference in differences. This workshop will cover foundational elements of modern practices of causal inference such as the potential outcomes model as well as discuss in detail the most common designs: regression discontinuity, instrumental variables, difference in differences, comparative case studies using synthetic control and if time permitting matching. It will be accompanied by efforts to introduce students to basic practices in programming as well as good research practices more generally.

Aims of the class:

1. To help students become more familiar with the field of causal inference
2. To empower students to apply research designs more competently to their own research
3. To direct students towards better programming practices so that they are better able to perform quantitative forms of research

**Group assignments**

We will devote 30 minutes for a potential outcomes assignment, 45 minutes for an RDD assignment, 45 minutes for an IV assignment, 45 minutes for a DiD assignment and 45 minutes for a synthetic control assignment.

**Software**

Most of our discussions will be in Stata. Though I have some familiarity with R, and have some code to share, most of my human capital is in Stata.

**Daily Structure**

This is a 3-day workshop. The goal of the workshop is for students to gain enough knowledge from the lectures and experience from the programming activities that they become confident and capable enough to implement and interpret these methods in their own work, as well as continue to learn this new material on their own after the workshop concludes. Each day lasts 8 hours with 4 hours of lecturing, 2 separate 75 minute “coding together” sessions, and the remainder are breaks and lunch. We will be strict about holding the 2.5 hours of coding and therefore may sacrifice on breaks if we are running behind.

**Schedule**

Day one – Introduction, Causality, Counterfactuals (8 hours)

* Hidden curriculum material
  + Programming
  + Organization
  + Workflow
  + Professional development
  + Mental health
* Potential outcomes
  + Randomization and selection bias
  + Fisher’s sharp null (i.e., randomization inference)
* Directed acyclic graphs

Day two: Regression discontinuity, Difference-in-differences (8 hours)

* Sharp regression discontinuity
* Introduction to difference in differences
  + Without covariates
  + Inverse probability weighting
  + Doubly robust
* Bias of TWFE with differential timing
* Weighted group-time ATT
  + Callaway and Sant’Anna
  + Sun and Abraham

Day three: Instrumental variables (8 hours)

* Synthetic control
* Instrumental variables
  + Some two step estimators (Wald, two sample IV, two stage least squares)
* Just identified
  + weak and strong instruments, bias and consistency of 2SLS
* Multiple instruments
  + LIML, jackknife, unbiased jackknife and double selection methods
* Heterogenous treatment effects and the local average treatment effect
* Leniency designs

Day four: Matching (8 hours)

* Subclassification
* Propensity scores
  + Weighting
  + Matching
* Nearest neighbor matching
* Coarsened exact matching
* Genetic matching