**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>

**Key programs (DiD – this will continue to expand):**

./programs/baker.do

./programs/baker\_cs.R

./programs/castle\_cs.R

./programs/five\_estimators\_examples.do

./programs/stacking\_castle.do

./programs/stacking\_castle\_broken.do

**Key subdirectory (synthetic control)**

./Texas

**Slides**

./Slides/codechella.pdf

**Assignments**

./Assignments/TBD

**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 is the co-director of the Health Services Research PhD program at Baylor and a former graduate of the University of Georgia (PhD, Economics) and the University of Tennessee at Knoxville (BA, Literature). 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 also the author of Causal Inference: the Mixtape published by Yale University Press in 2021. His research focus has traditionally been on topics in health and labor, including sex work, abortion policy, drug policy and mental healthcare. He has taught workshops on causal inference and difference-in-differences to universities and firms across the world such as Facebook, HP, University of Oxford, London School of Economics, University of Pennsylvania and many more.

**Workshop description:**

Our focus in this workshop is to learn to understand and to apply the “difference-in-differences” (DiD) methodology, a popular research design for causal inference the quantitative social sciences, as well as the synthetic control model. As some of these models are quite advanced, the goal is to provide “step downs” as much as possible through explainer style lecturing, Q&A and collaborative coding. The goal is that we will all reach basic competency and become conversant and literate about the methods and their implementation. The hope is that by the conclusion of the workshop, you will feel more comfortable about using some or all of these methods in your own research through greater understanding of the underlying econometrics and experience implementing them in Stata and/or R.

The workshop will cover 8 hours. Day one will cover issues with TWFE with and without covariates, with and without differential timing, with and without heterogeneity. We will conclude the day learning the syntax for implementation and then discussing those results together.

Day two will cover the imputation estimator and stacking. We will engage in a debugging exercise with the stacking do file, working together with problem solving to understand why the do file is flawed. If it goes quickly, then we will apply it to the baker dataset and discuss the results. After discussing imputation, I will show the different estimators against TWFE in a simulation to illustrate some of the main findings. If we have time, we will either discuss Sun and Abraham or synthetic control. My preference is the latter, because I think it’s worth a deep dive.

|  |  |  |
| --- | --- | --- |
| **DAY ONE (4 hours total)** | | |
| 30 min | Potential outcomes review; two by two without covariates | Mixtape chapter on potential outcomes and DiD |
| 60 min | Two by two with covariates | Abadie (2005); Sant’Anna and Zhao (2020) |
| 60 min | Bias of TWFE under differential timing | Goodman-Bacon (2021); simulation |
| 60 min | Manual aggregation methods: static and dynamic | Callaway and Sant’Anna (2020); simulation |
| 30 min | Coding lab | Implement CS with the Baker dataset after having studied the castle dataset syntax |

|  |  |  |
| --- | --- | --- |
| **DAY TWO (4 hours total)** | | |
| 60 min | Stacked regression | Minimum wages: Cengiz, et al. (2019); Clemens and Strain (2021) |
| 60 minutes | Coding Lab | Create a stacked dataset by debugging together a bad file. Apply the stacked dataset once it’s working to baker. |
| 60 min | Robust efficient imputation estimators | Borusyak, et al. (2021); Gardner (2021) |
| 30 min | Coding Lab | Application of BJS to baker dataset. |
| (if time) 30 min | Synthetic control | Abadie, Diamond and Hainmueller (2011) |