

# EITM Oslo 2024

## Strengthening the External Validity of Causal Designs

Nan Zhang  
n.zhang@uni-mannheim.de

### Course Description

At first glance, formal models and causal research designs seem like natural partners: while the former generates *ceteris paribus* predictions, the latter allow us to estimate causal quantities of interest holding all-else-constant. At the same time, modellers and empiricists may find themselves at loggerheads. Specifically, good theory should strive to make general predictions, but treatment effects are (almost) always locally-estimated. As a consequence, there often exists a mismatch between the targets of the formal model and the data used to evaluate these predictions. Strengthening the external validity of empirical approaches is therefore crucial for constructing stronger empirical tests of theory.

This course will introduce students to a framework for thinking through different types of external validity concerns. We will see how these concerns are manifest in both controlled experiments and natural experimental designs. Finally, students will learn methods for generalizing sample inferences to the target population.

### Course Requirements

You are expected to have completed the required readings (plus Supplementary Appendices) before class meets on Thursday and Friday. I will ask students to take the lead (either individually or in groups) in discussing the required readings. Your leadership includes providing an in-class summary of the readings and identifying discussion topics in the form of questions, reflections, or critiques.

Some examples of issues I would like to see covered in most discussions include:

- Is the empirical strategy appropriate for addressing the research question?
- What are the main strengths and limitations of the research design?
- If you do not find the empirical findings fully convincing, what pieces of evidence would have convinced you (i.e. what would you have done better)?
- What do the readings suggest about promising directions for future research?
- What are the lessons you draw for your own research?

But feel free to express any other ideas, questions, or comments that you may have!

## Part 1: Defining the Problem

In this introductory session, we will discuss different elements of external validity. I will also offer a primer on (or review of) the potential outcomes framework, instrumental variables, and regression discontinuity.

### Optional Background Readings:

Dunning, Thad. 2012. *Natural Experiments in the Social Sciences: A Design-based Approach*. Cambridge University Press.

Egami, Naoki, and Erin Hartman. 2023. “Elements of external validity: Framework, design, and analysis.” *American Political Science Review* 117 (3): 1070–1088.

## Part 2: Behavioral Experiments

Behavioral experiments are artificial in many ways. But when might they nonetheless be informative about theory? And when does the lack of external validity seriously limit what we can learn?

### Required Readings:

Druckman, James N, and Cindy D Kam. 2011. “Students as experimental participants.” *Cambridge handbook of experimental political science* 1:41–57.

Habyarimana, James, Macartan Humphreys, Daniel N Posner, and Jeremy M Weinstein. 2007. “Why does ethnic diversity undermine public goods provision?” *American political science review* 101 (4): 709–725.

Jerit, Jennifer, Jason Barabas, and Scott Clifford. 2013. “Comparing contemporaneous laboratory and field experiments on media effects.” *Public Opinion Quarterly* 77 (1): 256–282.

Martinangeli, Andrea FM, Marina Povitkina, Sverker Jagers, and Bo Rothstein. 2023. “Institutional quality causes generalized trust: experimental evidence on trusting under the shadow of doubt.” *American Journal of Political Science*.

## Part 3: Instrumental Variables

Instrumental variables have been used to two distinct settings: controlled experiments with non-compliance, and natural experiments where some but not all units end up receiving the treatment. We’ll use recent studies to illustrate how instrumental variables can work when applied to these settings and also see what can go wrong.

### Required Readings:

Bail, Christopher A, Lisa P Argyle, Taylor W Brown, John P Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. 2018. “Exposure to opposing views on social media can increase political polarization.” *Proceedings of the National Academy of Sciences* 115 (37): 9216–9221.

- Dinas, Elias, Konstantinos Matakos, Dimitrios Xeferis, and Dominik Hangartner. 2019. “Waking up the golden dawn: does exposure to the refugee crisis increase support for extreme-right parties?” *Political analysis* 27 (2): 244–254.
- Marbach, Moritz, and Dominik Hangartner. 2020. “Profiling compliers and noncompliers for instrumental variable analysis.” *Political Analysis* 28 (3): 435–444.
- Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. “Economic shocks and civil conflict: An instrumental variables approach.” *Journal of political Economy* 112 (4): 725–753.

## Part 4: Regression Discontinuity

Regression discontinuity analysis involves a natural experiment where treatment is assigned based on an arbitrary rule, like exceeding a threshold. While RDD is well-known for the extreme local-ness of estimated treatment effects, issues of compound treatments arise here as well.

### Required Readings:

- Lang, Valentin and Stephan Schneider. ”Immigration and Nationalism in the Long Run.” Link to Working Paper.
- Eggers, Andrew C. 2015. “Proportionality and turnout: Evidence from French municipalities.” *Comparative Political Studies* 48 (2): 135–167.
- Hainmueller, Jens, Dominik Hangartner, and Giuseppe Pirotta. 2017. “Catalyst or crown: Does naturalization promote the long-term social integration of immigrants?” *American Political Science Review* 111 (2): 256–276.
- Mo, Cecilia Hyunjung, and Katharine M Conn. 2018. “When do the advantaged see the disadvantages of others? A quasi-experimental study of national service.” *American Political Science Review* 112 (4): 721–741.

## Part 5: Designing Opportunities for Cumulative Learning

In an effort to assess the generalizability of treatment effects across populations and contexts, scholars (or teams of scholars) are increasingly conducting experiments around the same research questions in multiple locations. We will examine the metaketa initiative as a case-study in what one can learn from such endeavors.

**Required Readings:** Please read chapters 1-3, 11 (skim pp.327-347) and 13 in the following :

- Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D Hyde, Craig McIntosh, and Gareth Nellis. 2019. *Information, Accountability, and Cumulative Learning: Lessons from Metaketa I*. Cambridge University Press.