

# On Theory and Identification: When and Why We Need Theory for Causal Identification

Tara Slough\*

May 10, 2019

*Preliminary.*

## **Abstract**

What is the role for theory in identification-driven research designs? In this paper, I argue that not only is theory important for the interpretation of causal findings, but in many cases, theory is necessary for the identification of standard causal estimands (i.e., the ATE). In particular, I show that when empiricists study a sequence of post-treatment behavioral outcomes, post-treatment selection can prevent the identification of standard causal estimands, even when standard (empirical) identification assumptions hold. In these cases, articulation of a theory, or model of the world, that defines the post-treatment selection processes is necessary to define a set of identified causal estimands. Using a stylized example of crime, reporting, and recording, I illustrate how theory is necessary to reveal the set identified causal estimands, holding constant the research design. I then present this result more generally by considering the conditions under which the invocation of a theory is necessary for causal identification. I consider the implications for different research designs. This paper illustrates the need for theory in many identification-driven research designs.

---

\*Ph.D. Candidate, Columbia University and Predoctoral Fellow, University of California, Berkeley, [tls2145@columbia.edu](mailto:tls2145@columbia.edu). I thank Danny Hidalgo, Winston Lin, Fredrik Sävje and audiences at the Harvard Experiments Working Group and Yale Quantitative Methods Seminar for helpful comments. This project is supported in part by an NSF Graduate Research Fellowship, DGE-11-44155.