What Makes Treatment Effects Identifiable?
Characterizations and Estimators Beyond Unconfoundedness

Yang Cai   YANG.CAI@YALE.EDU
Alkis Kalavasis   ALKIS.KALAVASIS@YALE.EDU
Katerina Mamali   KATERINA.MAMALI@YALE.EDU
Anay Mehrotra   ANAYMEHROTRA1@GMAIL.COM
Manolis Zampetakis   MANOLIS.ZAMPETAKIS@YALE.EDU
Yale University

Editors: Nika Haghtalab and Ankur Moitra

Abstract

Most of the widely used estimators of the *average treatment effect* (ATE) in causal inference rely on the assumptions of *unconfoundedness* and *overlap*. Unconfoundedness requires that the observed covariates account for all correlations between the outcome and treatment. Overlap requires the existence of randomness in treatment decisions for all individuals. Nevertheless, many types of studies frequently violate unconfoundedness or overlap, for instance, observational studies with deterministic treatment decisions – popularly known as Regression Discontinuity designs – violate overlap.

In this paper, we initiate the study of general conditions that enable the *identification* of the average treatment effect, extending beyond unconfoundedness and overlap. In particular, following the paradigm of statistical learning theory, we provide an interpretable condition that is sufficient and nearly necessary for the identification of ATE. Moreover, this condition characterizes the identification of the *average treatment effect on the treated* (ATT) and can be used to characterize other treatment effects as well. To illustrate the utility of our condition, we present several well-studied scenarios where our condition is satisfied and, hence, we prove that ATE can be identified in regimes that prior works could not capture. For example, under mild assumptions on the data distributions, this holds for the models proposed by Tan (2006) and Rosenbaum (2002), and the Regression Discontinuity design model introduced by Thistlethwaite and Campbell (1960). For each of these scenarios, we also show that, under natural additional assumptions, ATE can be estimated from finite samples.

We believe these findings open new avenues for bridging learning-theoretic insights and causal inference methodologies, particularly in observational studies with complex treatment mechanisms.

Keywords: Causal Inference, Observational Studies, Hidden Biases, Regression Discontinuity Designs, Overlap, Unconfoundednes, Ignorability

---

Extended abstract. Full version appears as (Cai et al., 2025, v1).

References

Yang Cai, Alkis Kalavasis, Katerina Mamali, Anay Mehrotra, and Manolis Zampetakis. What makes treatment effects identifiable? characterizations and estimators beyond unconfoundedness, 2025. URL https://arxiv.org/abs/2506.04194.

Paul R. Rosenbaum. *Observational Studies*. Springer Series in Statistics. Springer, 2002. ISBN 9780387989679. URL https://books.google.com/books?id=K0OglGXtpGMC.

Zhiqiang Tan. A Distributional Approach for Causal Inference Using Propensity Scores. *Journal of the American Statistical Association*, 101(476):1619–1637, 2006. doi: 10.1198/016214506000000023. URL https://doi.org/10.1198/016214506000000023.

Donald L. Thistlethwaite and Donald T. Campbell. Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, 51(6):309–317, 1960. doi: 10.1037/h0044319.