RANDOMIZED CONTROLLED TRIALS

Called the "gold standard" because it costs a lot of money

Created by Sarah Miller (mille@umich.edu) **for BE 887**

Last edit April 6, 2022

MICHIGAN
ROSS SCHOOL OF BUSINESS

In a randomized controlled trial, the researcher randomly assigns some units to be treated and other units to be untreated. Because treatment is assigned at random, concern about selection into the treatment group is minimal.

WHEN TO USE

Requirement 1

The researcher controls treatment assignment, or is partnering with an organization that is willing to assign treatment at random

- Ideally the researcher can gather information about units prior to random assignment.

WHAT TO DO: THE BASICS

Step 1

Randomly assign some units to the treatment group and others to the control group

- Conduct a power analysis to determine the size of the sample you will need to detect effects. Power is maximized when half the sample is treated and half is not.
- Prior to randomly assigning treatment, post your pre-analysis plan on the AEA RCT Registry or the Open Science Framework.
- Verify that control and treatment groups are balanced across characteristics.
- Alternatively, assign treatment using a stratification mechanism that makes sure these groups are balanced.
  - E.g., if you are assigning half of the sample to treatment, assign one woman to the control group for each woman you assign to the treatment group.
  - This will assure you will be perfectly balanced on gender.

Step 2

Estimation

- Estimate a model of the form $y_i = \beta_0 + \beta_{RCT} D_i + \beta_1 X_i + \epsilon_i$, where $D_i = 1$ if i is assigned to the treatment group.
- If there is perfect compliance (i.e., no one in the control group is treated, and everyone in the treatment group is treated), then the coefficient β_{RCT} will give you the ATE and ATT.
- If there is imperfect compliance, you can estimate the LATE using treatment group assignment as an instrument for treatment receipt.
- Including covariates (X_i) measured prior to treatment can improve precision. Pre-treatment measures of the outcome may be particularly good for this purpose.
- If different strata have different treatment group assignment probabilities, you must include strata fixed effects.

INFERENCE

- If you are treating entire "clusters" of observations (e.g., entire schools or firms), but your data are at a lower level (e.g., students or workers), cluster your standard errors at the level of the treatment.
- Small number of clusters? No problem! Try methods in Cameron, Gelbach, and Miller 2008 or Hagemann 2019.
- If you are evaluating a large number of outcomes, consider adjusting your inference for the fact that you are testing multiple hypotheses (Anderson 2008).
- Want to test a "sharp" null hypothesis, that the treatment effect is 0 for all participants (rather than the effect is 0 on average)? Try randomization inference! See, e.g., Young 2019.

HOW THE PROS DO IT

- Ask other researchers what they think the outcome of the RCT will be, using `https://socialscienceprediction.org/`.
- How much external validity does your experiment have? Consider examining this using methods such as Kowalski 2019.

RATING

Difficulty
Fun
Validity



MAKE IT SIZZLE

- If you have the sample to support multiple treatment arms, try out variations of your treatment to better understand mechanisms.

SOURCES

References

- Hagemann, Andreas (2019). “Placebo inference on treatment effects when the number of clusters is small”. In: Journal of Econometrics 213.1.
- Kowalski, Amanda (2019). “How to Examine External Validity Within an Experiment”. In: NBER Working Paper 24834.
- Young, Alwin (2019). “Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results”. In: Quarterly Journal of Economics 134.
- Anderson, Michael (2008). “Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects”. In: JASA 103.484.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2008). “Bootstrap-Based Improvements for Inference with Clustered Errors”. In: The Review of Economics and Statistics 90.3.