Introduction to the Design and Analysis of Randomized Experiments
Class 5: Encouragment Designs, Instrumental Variables, Randomized Assignment but Not-Randomized Compliance

Jake Bowers

July 31, 2024

Overview and Review

Missing data in experiments

Non-compliance: Using Placebo Designs to estimate Causal effects when we do not control the dose

Learning about the CACE/LATE using randomization as an instrument

Open Discussion

Overview and Review

Today

0. Quiz and Questions
1. Missing data in experiments
2. Designs and assumptions to learn about the LATE/CACE: - When you can randomize something but not everything, what can you learn? (Answer, quite a lot under some assumptions either using randomization as an instrumental variable or using a placebo design.)
3. Open Discussions: Topics new and old. Things we didn't get around to discussing. Things that were confusing.

Note: You can download (and contribute to) course materials at https://github.com/bowers-illinois-edu/short_course_experiments

Hay recursos en español aqui: https://egap.github.io/theory_and_practice_of_field_experiments_spanish/

Lingering Questions?

Questions arising?

Quiz and Questions

- What are the three main ingredients of statistical power of a hypothesis test?
- In terms of those ingredients:
  - Why would a rank-based test statistic or a top-coded (or otherwise trimmed mean) based test statistic have more power than a simple difference of means?
  - Why would "covariance adjustment" improve power?
  - Why would block-randomization improve power?
  - Why might using an index of multiple measures of an outcome improve power?
- How might we make some informed guesses about the power of statistical tests given a research design?
- Why would a researcher publish an analysis plan before fielding an experiment? Why would academic journals start to require a "pre-analysis plan"?

Key points for this lecture

- Attrition (missing data on outcomes) can lead to bias and confounding and problems specifying the correct randomization distribution.
- Non-compliance can teach us about the effect of the "active ingredient" of the treatment.
  - We can learn about the causal effect of the active ingredient using randomization as an instrument
  - Or we can learn about the same causal effect using a placebo design (Nickerson, 2005),(Broockman and Kalla, 2016).

Missing data in experiments

Attrition (missing data on outcomes)

- Some units may have missing data on outcomes (= units attrit) when:
  - some respondents can't be found or refuse to participate in endline data collection.
  - some records are lost.
- This is a problem when treatment affects missingness.
  - For example, units in control may be less willing to answer survey questions.
  - For example, treatment may have caused units to migrate and cannot be reached
- If we analyze the data by dropping units with missing outcomes, then we are no longer comparing similar treatment and control groups.
  - If observed and/or unobserved covariates drive the process of missing outcomes, then treated and control groups no longer have the same probability of treatment.

What can we do?

- Check whether attrition rates are similar in treatment and control groups.
- Check whether treatment and control groups have similar covariate profiles.
- Do not drop observations that are missing outcome data from your analysis.
- When outcome data are missing we can sometimes **bound** our estimates of treatment effects.
  - Set the missing outcome data to the maximum value and then estimate a treatment effect
  - Set the missing outcome data to the minimum value and then estimate a treatment effect

What can we do?

- But the best approach is to try to anticipate and prevent attrition.
  - Blind people to their treatment status.
  - Promise to deliver the treatment to the control group after the research is completed.
  - Plan ex ante to reach all subjects at endline.
  - Budget for intensive follow-up with a random sample of attriters.

Missing data on covariates is not as problematic

- Missing **background covariates** (i.e.,variables for which values do not change as a result of treatment) for some observations is less problematic.
  - We can still learn about the causal effect of an experiment without those covariates, as we saw in the Hypothesis Testing and the Estimation modules.
  - We can also use the background covariate as planned by imputing for the missing values.
  - We can also condition on that missingness directly.

Missing data on treatment assignment or blocks or clusters is a big problem

This would cause very similar problem as would missing data on outcomes.

Summary: Missing data in experiments

- Missing data on covariates is not a problem: it should be balanced by randomization.
- Missing data on outcomes is a problem: it can remove the benefits of the randomization.
- Missing data on treatment assignment is a problem: it can remove the benefits of the randomization.

Given missing data on outcomes:

- Check to see if treatment predicts missingness
- Check to see if missingness patterns relate strongly to covariates.
- Report results using bounds (set all outcomes to max and min, report both results)
- Report results using sensitivity analysis (like an observational study (see Rosenbaum (2017))).

Non-compliance: Using Placebo Designs to estimate Causal effects when we do not control the dose

Non-compliance or Encouragement Designs I

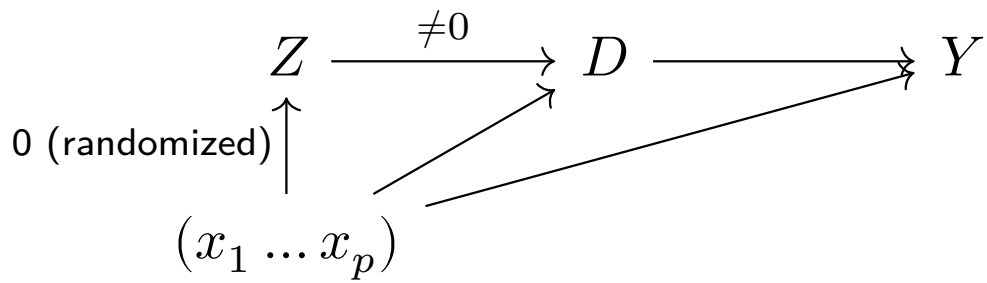
- Sometimes units assigned to treatment don't take it. They don't comply with their assignment.
  - If all units assigned to control do not take the treatment, but only some units assigned to treatment take the treatment, we have *one-sided non-compliance* (Gerber and Green, 2012, Chapter 5).
- The effect of **treatment assignment** (often called the Intent to Treat Effect or ITT) is not the same as the effect of **receiving a dose of the treatment**.
  - Not random who takes a dose of the treatment. Directly comparing people who take a dose with those who did not ("per protocol" or "as treated" comparisons) is not justified by randomization.

Non-compliance or Encouragement Designs I

- The effect of receiving the treatment is often called the "Local Average Treatment Effect" (LATE) or "Complier Average Causal Effect" (CACE) (same quantity).
  - "Local" refers to the idea that the effect only occurs on the people who take the treatment when assigned to treatment (the kinds of people who 'comply' with the intentions of the researcher).
- Two main approaches to estimation:
  1. Instrumental variable (in which random assignment may be a good instrument and often we estimate the CACE using 2SLS) (See Gerber and Green (2012) Chapter 5) and
  2. Placebo-controlled experiments (Nickerson (2005), example Broockman and Kalla (2016)). In which we estimate the CACE by comparing outcomes of treated and placebo groups.

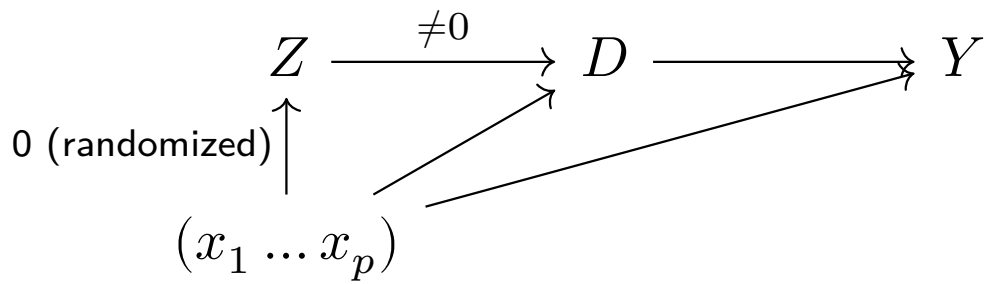
How to learn about the CACE/LATE using a Placebo

Imagine a door-to-door communication experiment where some houses are randomly assigned to receive a visit and some attitude is measured as an outcome later. Note that we now use Z (treatment assigned) and D (treatment taken) instead of just T.



- Z_i is random assignment to a visit ($Z_i = 1$) or not ($Z_i = 0$).
- $D_i(Z_i = 1) = 1$ means that person i would open the door to have a conversation when assigned a visit (it is a potential dose)
- $D_i(Z_i = 1) = 0$ means that person i would not open the door to have a conversation when assigned a visit.
- $x_1 \ldots x_p$ are background covariates — confounders of the $D_i \rightarrow Y_i$ relationship.

How to learn about the CACE/LATE using a Placebo



- Z_i is random assignment to a visit ($Z_i = 1$) or not ($Z_i = 0$).
- $D_i(Z_i = 1) = 1$ means that person i would open the door to have a conversation when assigned a visit (it is a potential dose)
- $D_i(Z_i = 1) = 0$ means that person i would not open the door to have a conversation when assigned a visit.
- $x_1 \ldots x_p$ are background covariates — confounders of the $D_i \rightarrow Y_i$ relationship.
- So, D_i is **an intermediate outcome:** $D_i = Z_i * D_i(1) + (1 - Z_i)D_i(0)$.
- $Y_i(Z_i = 1, D_i(Z_i = 1) = 1)$ is primary outcome when i is assigned to treatment **and** would open the door.

Defining causal effects

So now we have four potential outcomes per person.

- $Y_i(Z_i = 1, D_i = 1)$ is the potential outcome for people who were assigned a visit and who opened the door. ("Compliers" or "Always-takers")
- $Y_i(1, D_i = 0)$ is the potential outcome for people who were assigned a visit and who did not open the door. ("Never-takers" or "Defiers")
- $Y_i(0, D_i = 1)$ is the potential outcome for people who were not assigned a visit and who opened the door. ("Defiers" or "Always-takers")
- $Y_i(0, D_i = 0)$ is the potential outcome for people who were not assigned a visit and who did not open the door. ("Compliers" or "Never-takers")

Defining causal effects

The average causal effect of the "intention to visit people" is the ITT (averaging over whether or not people opened the door):

$$ITT = ITT_Y = \delta = \bar{Y}(Z = 1) - \bar{Y}(Z = 0). \quad (1)$$

The average causal effect of talking with someone at the door is the average treatment effect on compliers (the type of person who would open the door if visited) or the CACE:

$$CACE = \bar{Y}(Z_i = 1, D_i = 1 | D_i(1) = 1) - \bar{Y}(Z_i = 0, D_i = 0 | D_i(1) = 1) \quad (2)$$

This is the same as the Local Average Treatment Effect (LATE).

How to estimate the CACE using a Placebo Design?

We can estimate the average outcome among those assigned to treatment who opened their doors ($\bar{Y}(Z_i = 1, D_i = 1 | D_i(1) = 1)$) because there are no "always-takers" (i.e. people can only talk with a canvasser if a canvesser shows up).

$$\hat{\bar{Y}}_i(Z_i = 1, D_i = 1 | D_i(1) = 1) = \sum_i \frac{(Z_i D_i Y_i)}{\sum_i D_i} \tag{3}$$

Key here: **we use the design and context knowledge to claim that only compliers open the door when visited.**

How to estimate the CACE using a Placebo Design?

What about $\bar{Y}(Z_i = 0, D_i = 0 | D_i(1) = 1)$? How to learn about the outcome of people not assigned treatment but who would have opened their door if visited? (Hmm...We don't know whether they would have opened the door because they were not assigned a visit!)

What about a placebo? The Nickerson placebo-controlled design adds an arm to the experiment. For example, $Z_i = 2$, where:

- **the compliance should be the same as treatment (**$D_i(Z_i = 2) = D_i(Z_i = 1)$**)** and
- **outcome should be the same as control (**$Y_i(Z_i = 2) = Y_i(Z_i = 0)$**)**.

And where the $Z_i = 2$ is independent of all $x_1 \ldots x_p$ because of randomization. So people in $Z_i = 2$ are not systematically different from people in $Z_i = 1$ or $Z_i = 0$.

How to estimate the CACE using a Placebo Design?

This means that for:

$$CACE = \bar{Y}(Z_i = 1, D_i = 1 | D_i(1) = 1) - \bar{Y}(Z_i = 0, D_i = 0 | D_i(1) = 1) \quad (4)$$

We can estimate it with

$$\widehat{CACE} = \left(\sum_i \frac{(I(Z_i = 1) D_i Y_i)}{\sum_i I(Z_i = 1) D_i} \right) - \left(\sum_i \frac{(I(Z_i = 2) D_i Y_i)}{\sum_i I(Z_i = 2) D_i} \right) \quad (5)$$

Notice: we can check the assumptions:

- same compliance pattern (and same complier types between treatment and placebo): $D_i(Z_i = 2) = D_i(Z_i = 1)$
- same outcome pattern between placebo and control: $Y_i(Z_i = 2) = Y_i(Z_i = 0)$

Example from Broockman and Kalla using Nickerson

For example see that paper.

Summary

- Analyze as you randomized, even when you don't control the dose, to get the ITT (very common in policy contexts).
- The danger of per-protocol analysis (comparing based on D_i is an observational study, not a randomized experiment)
- If a placebo-controlled study doesn't work for you, you can use an instrumental variable approach (see Gerber and Green (2012) chapter 5)

Learning about the CACE/LATE using randomization as an instrument

Defining causal effects

We have four potential outcomes per person.

- $Y_i(Z_i = 1, D_i = 1)$ is the potential outcome for people who were assigned a visit and who opened the door. ("Compliers" or "Always-takers")
- $Y_i(1, D_i = 0)$ is the potential outcome for people who were assigned a visit and who did not open the door. ("Never-takers" or "Defiers")
- $Y_i(0, D_i = 1)$ is the potential outcome for people who were not assigned a visit and who opened the door. ("Defiers" or "Always-takers")
- $Y_i(0, D_i = 0)$ is the potential outcome for people who were not assigned a visit and who did not open the door. ("Compliers" or "Never-takers")

Defining causal effects II

We could also write $Y_i(Z_i = 0, D_i(Z_i = 1) = 1)$ for people who were not assigned a visit but who would have opened the door had they been assigned a visit etc.

In this case we can simplify our potential outcomes:

$Y_i(0, D_i(1) = 1) = Y_i(0, D_i(1) = 0) = Y_i(0, D_i(0) = 0)$ because your outcome is the same regardless of how you don't open the door because you were not assigned a visit.

Defining causal effects III

We can simplify the ways in which people get a dose of the treatment like so

- Y : outcome ($Y_i(Z_i = 1)$ for potential outcome to treatment for person i, fixed)
- X : covariate/baseline variable
- Z : treatment assignment ($Z_i = 1$ if assigned to a visit, $Z_i = 0$ if not assigned to a visit)
- D : treatment received ($D_i = 1$ if answered phone, $D_i = 0$ if person i did not answer the door) (using D here because $D_i = D_i(1)Z_i + D_i(0)(1 - Z_i)$)

Defining causal effects IV

We have two causal effects of Z: $Z \to Y$ (this effect is often known as δ, ITT, ITT_Y), and $Z \to D$ (Gerber and Green call this ITT_D).

And different types of people can react differently to the attempt to move the dose with the instrument.

		$Z = 1$	
		$D = 0$	$D = 1$
$Z = 0$	$D = 0$	Never taker	Complier
	$D = 1$	Defier	Always taker

Defining causal effects VI

The $ITT = ITT_Y = \delta = \bar{Y}(Z = 1) - \bar{Y}(Z = 0)$.

But, in this design, we can split $\bar{Y}(Z = 1) = \bar{Y}(1)$ into pieces: the outcome of those who answered the door (Compliers and Always-takers and Defiers). Write p_C for the proportion of compliers in the study.

$$\bar{Y}(1) = (\bar{Y}(1)|C)p_C + (\bar{Y}(1)|A)p_A + (\bar{Y}(1)|N)p_N + (\bar{Y}(1)|D)p_D. \quad (6)$$

And $\bar{Y}(0)$ is also split into pieces:

$$\bar{Y}(0) = (\bar{Y}(0)|C)p_C + (\bar{Y}(1)|A)p_A + (\bar{Y}(0)|N)p_N + (\bar{y}_0|D)p_D. \quad (7)$$

Defining causal effects VII

So, the ITT itself is a combination of the effects of Z on Y within these different groups (imagine substituting in and then re-arranging so that we have a set of ITTs, one for each type of subject). But, we can still estimate it because we have unbiased estimators of $\bar{Y}(1)$ and $\bar{y}_0$ within each type.

Learning about the ITT I

First, let's learn about the effect of the policy itself. To write down the ITT, we do not need to consider all of the types above. We have no defiers ($p_D = 0$) and we know the ITT for both Always-takers and Never-takers is 0.

$$\bar{Y}(1) = (\bar{Y}(1)|C)p_C + (\bar{Y}(1)|A)p_A + (\bar{Y}(1)|N)p_N \tag{8}$$

$$\bar{Y}(0) = (\bar{Y}(0)|C)p_C + (\bar{Y}(0)|A)p_A + (\bar{Y}(0)|N)p_N \tag{9}$$

Learning about the ITT II

First, let's learn about the effect of the policy itself. To write down the ITT, we do not need to consider all of the types above. We have no defiers ($p_D = 0$) and we know the ITT for both Always-takers and Never-takers is 0.

$$
\begin{aligned}
ITT =& \bar{Y}(1) - \bar{Y}(0) && (10)\\
=& ((\bar{Y}(1)|C)p_C + (\bar{Y}(1)|A)p_A + (\bar{Y}(1)|N)p_N) - && (11)\\
& ((\bar{Y}(0)|C)p_C + (\bar{Y}(0)|A)p_A + (\bar{Y}(0)|N)p_N) && (12)
\end{aligned}
$$

collecting each type together — to have an ITT for each type

$$
\begin{aligned}
=& ((\bar{Y}(1)|C)p_C - (\bar{Y}(0)|C)p_C) + ((\bar{Y}(1)|A)p_A - (\bar{Y}(1)|A)p_A) + && (13)\\
& ((\bar{Y}(1)|N)p_N - (\bar{Y}(0)|N)p_N) && (14)\\
=& \left((\bar{Y}(1)|C) - (\bar{Y}(0)|C)\right)p_C + && (15)\\
& \left((\bar{Y}(1)|A) - (\bar{Y}(0)|A)\right)p_A + \left((\bar{Y}(1)|N) - (\bar{Y}(0)|N)\right)p_N && (16)
\end{aligned}
$$

Learning about the ITT III

$$
\begin{aligned}
ITT =& \bar{Y}(1) - \bar{Y}(0) && (17)\\
=& ((\bar{Y}(1)|C)p_C + (\bar{Y}(1)|A)p_A + (\bar{Y}(1)|N)p_N) - && (18)\\
& ((\bar{Y}(0)|C)p_C + (\bar{Y}(0)|A)p_A + (\bar{Y}(0)|N)p_N) && (19)\\
=& ((\bar{Y}(1)|C)p_C - (\bar{Y}(0)|C)p_C) + ((\bar{Y}(1)|A)p_A - (\bar{Y}(1)|A)p_A) + && (20)\\
& ((\bar{Y}(1)|N)p_N - (\bar{Y}(0)|N)p_N) && (21)\\
=& ((\bar{Y}(1)|C) - (\bar{Y}(0)|C))p_C + ((\bar{Y}(1)|A) - (\bar{Y}(0)|A))p_A + && (22)\\
& ((\bar{Y}(1)|N) - (\bar{Y}(0)|N))p_N && (23)
\end{aligned}
$$

Learning about the ITT IV

And, if the effect of the dose can only occur for those who open the door, and you can only open the door when assigned to do so then:

$$((\bar{Y}(1)|A) - (\bar{Y}(0)|A))p_A = 0 \text{ and } ((\bar{Y}(1)|N) - (\bar{Y}(0)|N))p_N = 0 \quad (24)$$

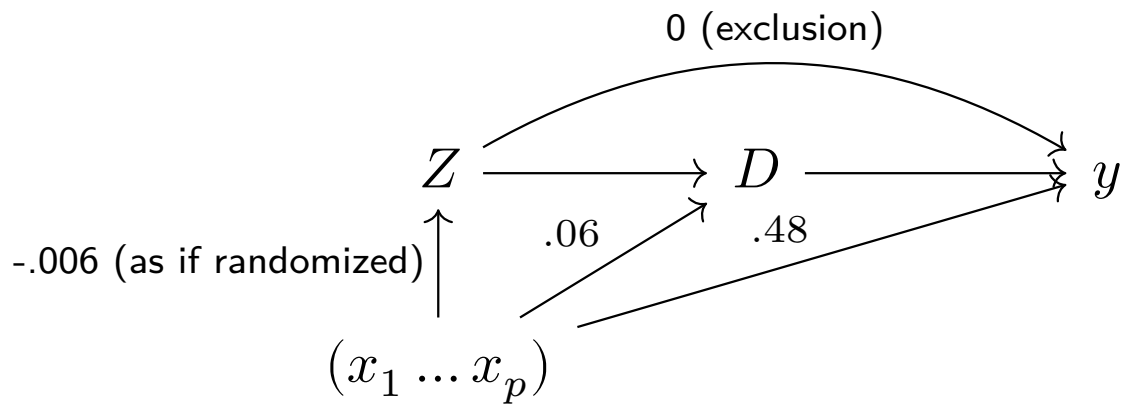
And

$$ITT = ((\bar{Y}(1)|C) - (\bar{Y}(0)|C))p_C = (CACE)p_C. \quad (25)$$

The complier average causal effect I

We would also like to learn about the causal effect of answering the door and having the conversation, the theoretically interesting effect.

But this comparison is confounded by x: a simple $\bar{Y}|D=1 - \bar{Y}|D=0$ comparison tells us about differences in the outcome due to x in addition to the difference caused by D.



But we just saw that, in this design, and with these assumptions (including a SUTVA assumption) that $ITT = ((\bar{Y}(1)|C) - (\bar{Y}(0)|C))p_C = (CACE)p_C$, so we can define $CACE = ITT/p_C$.

How to calculate the ITT and CACE/LATE I

Some example data (where we know all potential outcomes):

X	u	type	Z	pZ	DZ1	YD0Z0	YD1Z0	YD0Z1	YD1Z1	D	Y
1	-0.24	Complier	0	0	1	-0.24	0.25	-0.24	0.25	0	-0.24
2	-1.03	Complier	1	0	1	-1.03	-0.54	-1.03	-0.54	1	-0.54
3	-0.71	Complier	0	0	1	-0.71	-0.23	-0.71	-0.23	0	-0.71
3	0.26	Always-Taker	0	1	1	0.26	0.74	0.26	0.74	1	0.74
1	-0.25	Always-Taker	1	1	1	-0.25	0.24	-0.25	0.24	1	0.24

How to calculate the ITT and CACE/LATE II

The ITT and CACE (the parts)

```
itt_y <- difference_in_means(Y ~ Z, data = dat0)
itt_y
```

```
## Design:  Standard
##    Estimate Std. Error  t value  Pr(>|t|)    CI Lower CI Upper
## Z 0.2938767  0.1964104 1.496238 0.1378134 -0.09590562 0.683659 97.
```

```
itt_d <- difference_in_means(D ~ Z, data = dat0)
itt_d
```

```
## Design:  Standard
##   Estimate Std. Error  t value     Pr(>|t|)  CI Lower  CI Upper
## Z      0.8 0.06006799 13.31824 1.127737e-22 0.6806015 0.9193985 86
```

How to calculate the ITT and CACE/LATE III

All together:[1]

```
cace_est <- iv_robust(Y ~ D | Z, data = dat0)
cace_est
```

```
##                Estimate Std. Error    t value Pr(>|t|)   CI Lower
## (Intercept) -0.04944467  0.1453309 -0.3402213 0.734418 -0.3378491
## D            0.36734587  0.2423256  1.5159186 0.132759 -0.1135414
```

```
## Notice same as below:
coef(itt_y)["Z"] / coef(itt_d)["Z"]
```

```
## [1] 0.3673459
```

[1] works when $Z \to D$ is not weak see Imbens and Rosenbaum (2005) for a cautionary tale

Summary of Instrumental Variable based approach to the CACE/LATE

- Analyze as you randomized, even when you don't control the dose
- If you randomize treatment **assignment**, you can estimate the average effect of the dose on those who comply under the five assumptions that we made above:
  - Z is randomized (so, breaking relationships with all x)
  - SUTVA
  - Z has an effect on D ("non-zero proportion of compliers")
  - No defiers
  - Z only affects Y through D ("exclusion restriction")
- The danger of per-protocol analysis (Comparing outcomes based on D returns you to the world of observational studies).

Open Discussion

Open Discussion

Your questions and thoughts

References I

- Broockman, David and Joshua Kalla (2016). "Durably reducing transphobia: A field experiment on door-to-door canvassing". In: *Science* 352.6282, pp. 220–224.
- Gerber, Alan S. and Donald P. Green (2012). *Field Experiments: Design, Analysis, and Interpretation*. New York, NY: W. W. Norton & Company.
- Imbens, Guido W. and Paul R. Rosenbaum (2005). "Robust, accurate confidence intervals with a weak instrument: quarter of birth and education". In: *Journal of the Royal Statistical Society Series A* 168.1, pp. 109–126.
- Nickerson, D.W. (2005). "Scalable protocols offer efficient design for field experiments". In: *Political Analysis* 13.3, p. 233.
- Rosenbaum, Paul R. (2017). *Observation and Experiment: An Introduction to Causal Inference*. Harvard University Press.