La implementación de experimentos de encuesta

Santiago López-Cariboni, Universidad Católica del Uruguay

July 26 2019

Sampling

How do we find participants?

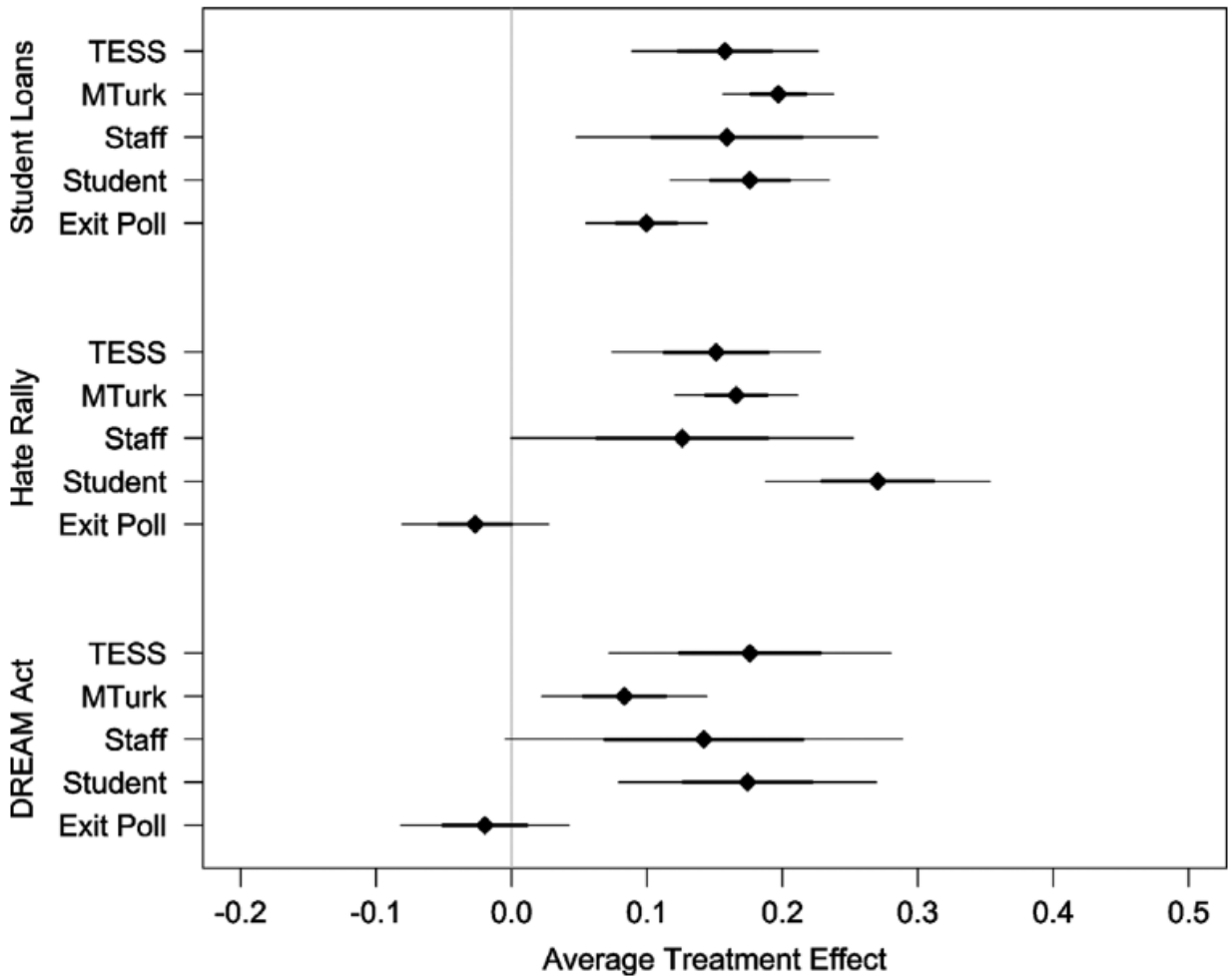
- Volunteers
  - Volunteer Science
  - In-house subject pool
- Online recruitment
  - Facebook
- Paid crowdworkers
  - Prolific Academic
  - Mechanical Turk
  - Crowdflower
- "Representative" samples
  - Big players: YouGov, TNS, Gallup, Nielsen, GfK
  - Others: Kantar, SSI, Lucid

Generalizability of Survey Experiments[1]

Convenient samples Vs population based-experiments.

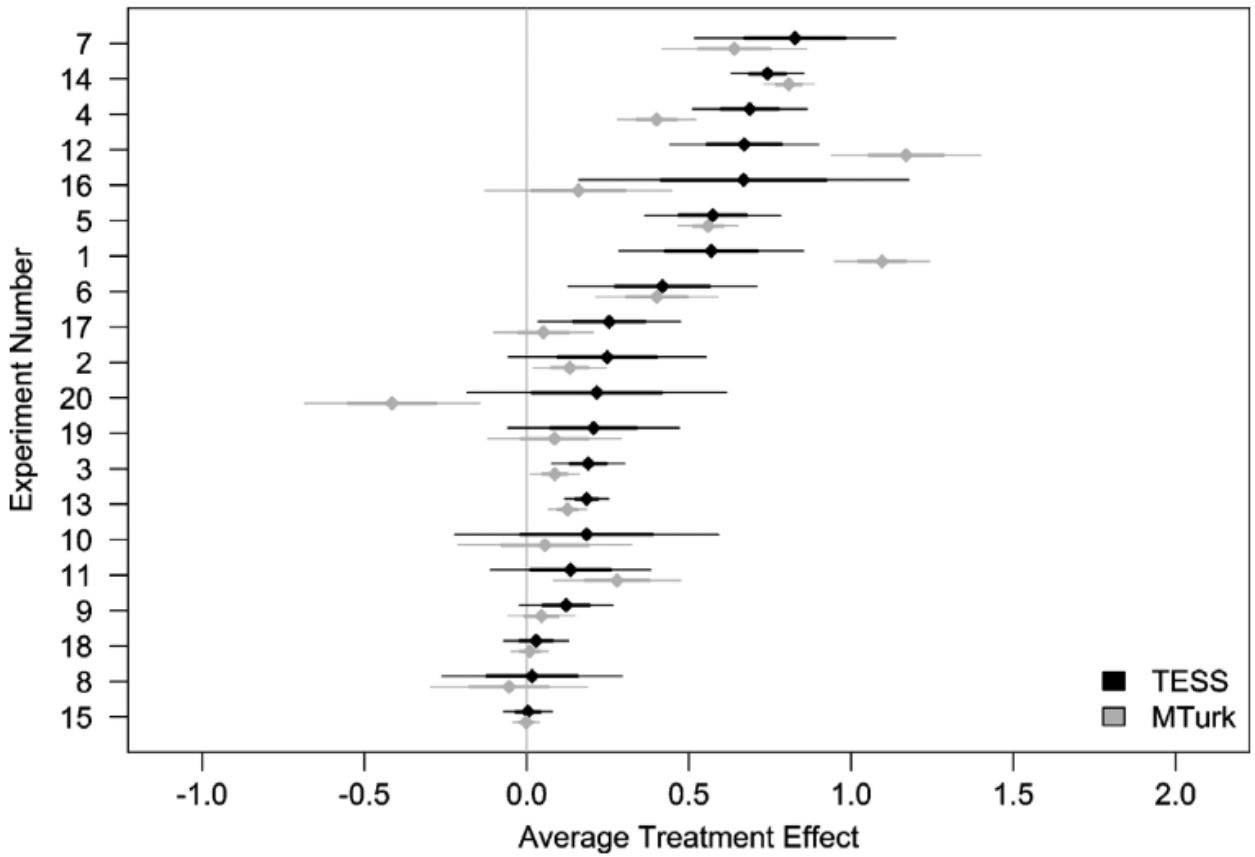
[1] Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman, and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." Journal of Experimental Political Science 2 (2): 109–38.

Generalizability of Survey Experiments[2]



[2]Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman, and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." Journal of Experimental Political Science 2 (2): 109–38.

Generalizability of Survey Experiments[3]



[3]Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman, and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." Journal of Experimental Political Science 2 (2): 109–38.

Attention, Satisficing, and Noncompliance

Dealing with misbehavior

Wow we handle or analyze survey-experimental data where we think participants misbehaved.

This falls into a couple of broad categories:

- Noncompliance
- Inattention
- Survey Satisficing

Dealing with misbehavior

How should we deal with respondents that appear to not be paying attention, not "taking" the treatment, or not responding to outcome measures?

1. Keep them
2. Throw them away

Best Practice: Pre-Analysis Protocol

- Excluding respondents based on survey behavior is one of the easiest ways to "p-hack" an experimental dataset
  - Inattention, satisficing, etc. will tend to reduce the size of the SATE
- So regardless of how you handle these respondents, these should be decisions that are made *pre-analysis*

When are you excluding participants?

- Pre-Treatment
  - Satisficing behaviors
  - Inattention
  - Covariate-based selection
  - Pretreated
- Post-Treatment
  - Speeding on treatment
  - "Failing" a manipulation check
  - Drop-off

Pre-Treatment Exclusion

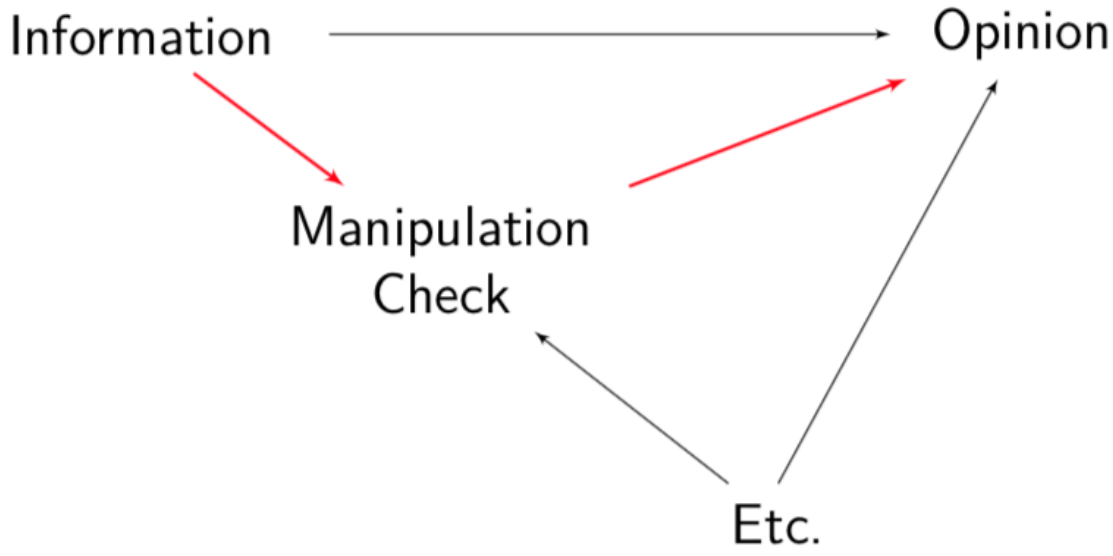
This is totally fine from a causal inference perspective

- Advantages:
- Focused on engaged respondents
- Likely increase impact of treatment
- Disadvantages: Changing definition of sample (and thus population)

Post-Treatment Exclusion

- This is much more problematic because it involves controlling for a post-treatment variable

Post-Treatment Exclusion



Risk that estimate of β_1 is diminished because effect is being carried through the manipulation check. Introduction of "collider bias" wherein values of the manipulation check are affected by other factors.

Post-Treatment Exclusion

- Any post-treatment exclusion is problematic and should be avoided
- Can estimate a LATE
  - Interpretation: Effect of manipulation check among those whose value of the check can be changed by the treatment manipulation
- Non-response or attrition is the same as researcher-imposed exclusion
  - Not problematic if MCAR
  - Nothing really to be done if caused by treatment

Apparent Satisficing

- Some common measures:
  - "Straightlining"
  - Non-differentiation
  - Acquiescence
  - Nonresponse
  - DK responding
  - Speeding
- Difficult to detect and distinguish from "real" responses

Metadata

- Timing
  - Some survey tools will allow you to time page
  - Make a prior rules about dropping participants for speeding
- Mousetracking or eyetracking
  - Mousetracking is unobtrusive
  - Eyetracking requires participants opt-in
- Record focus/blur browser events

Direct Measures

- How closely have you been paying attention to what the questions on this survey actually mean?
- While taking this survey, did you engage in any of the following behaviors? Please check all that apply.
  - Use your mobile phone
  - Browse the internet
  - ...

Attention questions

Do you agree or disagree with the decision to send British forces to fight ISIL in Syria? We would like to know if you are reading the questions on this survey. If you are reading carefully, please ignore this question, do not select any answer below, and click "next" to proceed with the survey.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

Treatment Noncompliance

Definition:

"when subjects who were assigned to receive the treatment go untreated or when subjects assigned to the control group are treated"[4]

Several strategies:

- "As treated" analysis
- "Intention to treat" analysis
- Estimate a LATE

[4] Gerber & Green. 2012. Field Experiments, p.132.

Analyzing Noncompliance

If noncompliance only occurs in one group, it is asymmetric or one-sided

We can ignore non-compliance and analyze the "intention to treat" effect, which will underestimate our effects because some people were not treated as assigned: $ITT = \bar{Y}_1 - \bar{Y}_0$

We can use "instrumental variables" to estimate the "local average treatment effect" (LATE) for those that complied with treatment: $LATE = \frac{ITT}{\%ofCompliant}$

Local Average Treatment Effect

- IV estimate is local to the variation in X that is due to variation in D
- This matters if effects are heterogeneous
- LATE is effect for those who comply
- Four subpopulations:
  - Compliers: $X = 1$ only if $D = 1$
  - Always-takers: $X = 1$ regardless of D
  - Never-takers: $X = 0$ regardless of D
  - Defiers: $X = 1$ only if $D = 0$

Regression Estimation

Regression Adjustment in Experiments:

- Recall the general advice that we do not need covariates in the regression to "control" for omitted variables (because there are none)
- Including covariates can reduce variance of our SATE by explaining more of the variation in Y

Scenario

Imagine two regression models. Which is correct?

- Mean-difference of SATE is not significant
- Mean-difference of SATE is with control variables

This is a small-sample dynamic, so make these decisions pre-analysis!

Treatment-Covariate Interactions

- The regression paradigm allows us to estimate CATEs using interaction terms
  - X is an indicator for treatment
  - M is an indicator for possible moderator
- SATE: $Y = \beta_0 + \beta_1 X + \epsilon$
- CATEs:

$$Y = \beta_0 + \beta_1 X + \beta_2 M + \beta_3 X \times M + \epsilon$$

  - Homogeneity: $\beta_3 = 0$
  - Heterogeneity: $\beta_3 \neq 0$