

Lab 9: Field Experiments II

Due: Monday, March 18, 11:59 PM

Name: Your name here

Mac ID: The first half of your Mac email address

1 Pre-post design

[Diaz and Rossiter \(2023\)](#) argue that, while implementing pre-post designs is generally a good idea, doing so sometimes entails sacrificing some of your sample size, which may offset the precision gains. Usually, this happens when research budgets are fixed, so that the extra resources that go into measuring pre-treatment outcomes come at the expense of the total sample size.

To illustrate this point, let's simulate a design that compares the standard design and the pre-post design, but implementing the pre-post design means losing a proportion of the observations.

The following code does this, but notice that we are missing the step of putting all parts of the model together.

```
drop = 0.2 # proportion dropped in pre-post design

# M

pp_model = declare_model(
```

```

N = 500,
U = rnorm(N)/2,
# time-specific unobserved factors
e1 = rnorm(N),
e2 = rnorm(N),
# R indicates if unit stays in sample,
R = rbinom(N, 1, prob = 1 - drop),
Y1 = U + e1, # outcome before treatment,
potential_outcomes(Y2 ~ 0.3*Z + Y1 + e2)
)

# I
pp_inq = declare_inquiry(
  ate = mean(Y2_Z_1 - Y2_Z_0)
)

# D
pp_assign = declare_assignment(
  Z = complete_ra(N)
)

pp_measure = declare_measurement(
  Y2 = reveal_outcomes(Y2 ~ Z)
)

# A
## estimate standard experiment

```

```

## with regression to make
## comparison fair
pp_standard = declare_estimator(
    Y2 ~ Z,
    .method = lm_robust,
    inquiry = "ate",
    label = "Standard"
)

pp_prepost = declare_estimator(
    Y2 ~ Z + Y1,
    .method = lm_robust,
    subset = R == 1,
    inquiry = "ate",
    label = "Pre-post"
)

```

Task 1

Compare the performance of the standard experiment and the pre-post experiment in terms of bias, RMSE. Which design would you recommend and why?

Does your recommendation change if the pre-post experiment now sacrifices half of the sample? Explain.

Does your recommendation change if the pre-post experiment now loses more than half of the sample? Explain.

2 Block randomization

A famous textbook on field experiments suggests that the biggest downside of conducting a block-randomized experiment is inadvertently analyzing them as if they assigned conditions through simple or complete randomization (Gerber and Green 2012).

How bad could this problem be? The following code simulates a block-randomized experiment and compares two answer strategies: A **naive** estimator and a **block** estimator. Once again, we are skipping the step of combining the different elements of the design into one.

```
# True ATEs per block
tau = c(4, 2, 0)

# M
b_model = declare_model(
  block = add_level(N = 3,
                    prob = c(0.5, 0.7, 0.9),
                    tau = tau),
  indiv = add_level(N = 100,
                    U = rnorm(N),
                    Y_Z_0 = U,
                    Y_Z_1 = U + tau)
)

# I
b_inq = declare_inquiry(
  ate = mean(Y_Z_1 - Y_Z_0)
)
```

```

# D
b_assign = declare_assignment(
  Z = block_ra(blocks = block,
               block_prob = c(.5, .7, .9)),
  Z_cond_prob = obtain_condition_probabilities(
    assignment = Z,
    blocks = block,
    block_prob = c(.5, .7, .9))
)

b_measure = declare_measurement(
  Y = reveal_outcomes(Y ~ Z)
)

# A
b_naive = declare_estimator(
  Y ~ Z,
  .method = difference_in_means,
  inquiry = "ate",
  label = "Naive"
)

b_block = declare_estimator(
  Y ~ Z,
  blocks = block,
  .method = difference_in_means,

```

```
inquiry = "ate",  
label = "Block"  
)
```

i Task 2

Compare both answer strategies in terms of bias, RMSE, and power. Which one would you recommend and why?

What happens if you pick three new values for the true ATEs per block stored in `tau`? Explain.

What happens if the three ATEs per block have the same value? Explain.

3 Answers

Remember to use `set.seed()` and to show the code and output for every step!

3.1 Task 1

Work on your answers here.

3.2 Task 2

Work on your answers here.