

Problem Set 4

Research Design for Causal Inference

Due: April 28, 2015

Part I – Project Star (redux)

Here we return to the Project STAR data that you started analyzing in Problem Set 2. I recommend revisiting the description of the study and variables presented there as well as the R code you used to perform the analysis before you try to answer the questions below.

Question 1 – Estimate treatment effects and conduct hypothesis tests

Part a Estimate the treatment effect using a difference-in-means estimator by conducting a t-test.

Part b Estimate the treatment effect using a linear regression model (OLS) with no covariates.

Part c Estimate the treatment effect using a linear regression model with any pre-treatment covariates you believe may be relevant.

Part d Are your estimates of treatment effects in Parts a, b and c different? If they are different, explain (with words and/or equations) why such differences might occur.

Part e Are there any reasons to prefer any of the estimators (difference-in-means; OLS without covariate adjustment; OLS with covariate adjustment) over the others?

Question 2 – Interpret the findings

What do you conclude about the effect of this intervention on the basis of these analyses (you may also refer to the analyses you reported in Problem Set 2)?

Bonus – Test for covariate imbalance using an F statistic

Recall how you used Chi-square tests to check covariate balance in Problem Set 2. Now, use the method described by Gerber & Green in *FEDAI*, Section 4.3 to check for covariate imbalance by regressing treatment assignment on the covariates and then creating a sampling distribution of F -statistics for the models estimated with and without covariates.¹

R hint: You will probably want to write a for-loop or a function to estimate regression models for the sampling (randomization) distribution with and without covariates and collect the sum of squared residuals (SSR) for each model. Finding the residuals for a linear model is pretty straightforward. For example, if you estimate and store the results of a linear model like this:

```
my.lm <- lm(y ~ x, data = d)
```

You can access the residuals by calling `my.lm$residuals`.

Part II – Gerber & Green Chapter 3 exercises

Excercise 7

Exercise 9 (Only parts a-c!)

Part III – Wantchekon revisited

For this question you'll use randomization inference to revisit the findings from Leonard Wantchekon's (2003) field experiment on the effect of clientelist vs. public policy framing for campaign messages on voter turnout in Benin.

The dataset is available either a [csv](#) or an [RData](#) file from the course website.

Recall that the structure of the experiment was *block* randomization. Villages were divided into groups of 3 based on geography and treatment status was randomized within the 8 groups of 3. The outcome variable is the vote share of the candidate participating in the experiment. The only covariate we'll use here is the number of registered voters. In the dataset, `block` indicates block group, `reg.voters` is the registered voters covariate, `vote.pop` is the outcome variable, `treat` is a variable indicating treatment status. The following questions will ask you to analyze the differences between the Clientelist and Public Policy conditions.

Part a Estimate the effect the clientelist message compared to the public policy message, using a regression estimator. For the purpose of comparison, do not include the covariate

¹ Please note that you should always check covariate balance *before* you conduct hypothesis tests, but since the purpose of this Question is to revisit your old tests (and it involves more complicated computation than Question 1), I've put them in the reverse order here.

`reg.vote`. Note that the block structure of this experiment will affect how you calculate these quantities. For the regression estimate, you can account for the blocks by including block level dummy variables (R users: include as `.factor(block)`) in your regression equation.²

Part b Now calculate an estimate of the treatment effect using the difference-in-means as your test statistic. Remember that the block structure of the experiment affects how you calculate this! (see *FEDAI* for details). To help you complete Part c, you'll want to write a function that does this in three steps:

1. Calculate the mean outcome within blocks for each experimental condition. R users: the `tapply()` function will help you do this. For both the Clientelist and Public Policy conditions, you'll want to run `tapply(outcome, block, mean)`.
2. Calculate weights that you will apply to the mean values for each block that you gathered in Step 1. Every weight, w_b , for every block, b , is equal to the proportion of the sample included in that block. Formally, $w_b = \frac{n_b}{N}$, where n_b is the size of the sample block and N is the total sample size.
3. Calculate the aggregate difference of means (your test statistic) estimate by multiplying the difference of means within each block, b by their respective weights, w_b and then taking the difference of these weighted estimates.

Part c Use randomization inference to test your difference of means estimate from Part b under the sharp null hypothesis of no effects. You can do this using the function described in Part b above and incorporating it into the randomization inference code from the previous problem set. You can also complete this using the "RI" package (I'll try to generate solution code for both). Note that you don't need to enumerate every possible treatment assignment to create the randomization distribution, but can sample a large number of draws (e.g., 5000) instead.

Part d What do you conclude about the effect of clientelist-oriented campaign messages versus those oriented towards public policy on voter turnout in Benin based on these results? How do your findings compare with Wantchekon's published findings?

Bonus Plot the randomization distribution using a kernel density plot and the true test statistic as an overlaid a vertical line. (R users can overlay the vertical line using something like: `abline(v=test.statistic, col='red', lwd=2)`).

²Note that this approach would produce biased estimates if the blocks were different sizes! See *FEDAI* exercise 4.9 for an illustration of this.

Part IV – Key concepts for next class

- The differences between a Gerber & Green’s ideas of a research proposal, preanalysis plan (planning document), report, and article (as well as the corresponding checklists)
- Researcher degrees of freedom in a “garden of forking paths.”
- Why multiple comparisons might be a problem.
- P-hacking and “fishing” (and why Gelman & Loken regret both terms).
- Internal validity vs. external validity (generalizability).
- Pre-registration of study protocols.
- Replication.