# Problem Set #4

*Experiment Design: Alex & Daniel*

```r
# load packages
library(foreign)
```

## FE exercise 5.2

a. Make up a hypothetical schedule of potential outcomes for three Compliers and three Never-Takers where the ITT is positive but the individual causal effect is negative. (*Note: By individual causal effect, I mean the comparison of $y_i(1) - y_i(0)$). This would not generally be possible without a science table.* If you want, you could make this data and demonstrate to yourself that although you can create an ITT and causal effect that are different signs with the science table, that you cannot create an ITT and CACE that are different signs on observed data.)
b. Suppose that an experiment were conducted on your pool of subjects. In what ways would the estimated CACE be informative or misleading?
c. **In addition, please also answer this question**: Which population is more relevant to study for future decision making: the set of Compliers, or the set of Compliers plus Never-Takers? Why?

## FE exercise 5.6

Suppose that a researcher hires a group of canvassers to contact a set of 1,000 voters randomly assigned to a treatment group. When the canvassing effort concludes, the canvassers report that they successfully contacted 500 voters in the treatment group, but the truth is that they only contacted 250. When voter turnout rates are tabulated for the treatment and control groups, it turns out that 400 of the 1,000 subjects in the treatment group voted, as compared to 700 of the 2,000 subjects in the control group (none of whom were contacted).

a. If you believed that 500 subjects were actually contacted, what would your estimate of the CACE be?
b. Suppose you learned that only 250 subjects were actually treated. What would your estimate of the CACE be?
c. Do the canvassers' exaggerated reports make their efforts seem more or less effective? Define effectiveness either in terms of the ITT or CACE. Why does the definition matter?

## FE exercise 5.10

Guan and Green report the results of a canvassing experiment conduced in Beijing on the eve of a local election. Students on the campus of Peking University were randomly assigned to treatment or control groups. Canvassers attempted to contact students in their dorm rooms and encourage them to vote. No contact with the control group was attempted. Of the 2,688 students assigned to the treatment group, 2,380 were contacted. A total of 2,152 students in the treatment group voted; of the 1,334 students assigned to the control group, 892 voted. One aspect of this experiment threatens to violate the exclusion restriction. At every dorm room they visited, even those where no one answered, canvassers left a leaflet encouraging students to vote.

```r
library(foreign)
d <- read.dta("./data/Guan_Green_CPS_2006.dta")
head(d)
```

```
##   turnout contact  dormid treat2
## 1       0       0 1010101      0
## 2       0       0 1010101      0
## 3       0       0 1010101      0
## 4       0       0 1010102      0
## 5       0       0 1010102      0
## 6       0       1 1010103      1
```

a. Using the data set from the book's website, estimate the ITT. First, estimate the ITT using the difference in two-group means. Then, estimate the ITT using a linear regression on the appropriate subset of data. *Heads up: There are two NAs in the data frame. Just na.omit to remove these rows.*

b. Use randomization inference to test the sharp null hypothesis that the ITT is zero for all observations, taking into account the fact that random assignment was clustered by dorm room. Interpret your results.

c. Assume that the leaflet had no effect on turnout. Estimate the CACE. Do this in two ways: First, estimate the CACE using means. Second, use some form of linear model to estimate this as well. If you use a 2SLS, then report the standard errors and draw inference about whether the leaflet had any causal effect among compliers.

d. *SKIP*

e. *SKIP*

f. *SKIP*

# FE exercise 5.11

Nickerson describes a voter mobilization experiment in which subjects were randomly assigned to one of three conditions: a baseline group (no contact was attempted); a treatment group (canvassers attempted to deliver an encouragement to vote); and a placebo group (canvassers attempted to deliver an encouragement to recycle). Based on the results in the table below answer the following questions

| Treatment Assignment | Treated ? | N | Turnout |
|---|---|---|---|
| Baseline | No | 2572 | 31.22% |
| Treatment | Yes | 486 | 39.09% |
| Treatment | No | 2086 | 32.74% |
| Placebo | Yes | 470 | 29.79% |
| Placebo | No | 2109 | 32.15% |

**First** Use the information to make a table that has a full recovery of this data. That is, make a `data.frame` or a `data.table` that will have as many rows a there are observations in this data, and that would fully reproduce the table above. (*Yes, this might seem a little trivial, but this is the sort of "data thinking" that we think is important.*)

a. We are rewriting part (a) as follows: "Estimate the proportion of Compliers by using the data on the Treatment group. Then compute a second estimate of the proportion of Compliers by using the data on the Placebo group. Are these sample proportions statistically significantly different from each other? Explain why you would not expect them to be different, given the experimental design." (Hint: ITT_D means "the average effect of the treatment on the dosage of the treatment." I.E., it's the contact rate $\alpha$ in the async).

b. Do the data suggest that Never Takers in the treatment and placebo groups have the same rate of turnout? Is this comparison informative?

c. Estimate the CACE of receiving the placebo. Is this estimate consistent with the substantive assumption that the placebo has no effect on turnout?

d. Estimate the CACE of receiving the treatment using two different methods. First, use the conventional method of dividing the ITT by the ITT_{D}.

e. Then, second, compare the turnout rates among the Compliers in both the treatment and placebo groups. Interpret the results.

# FE exercise 8.10

A doctoral student conducted an experiment in which she randomly varied whether she ran or walked 40 minutes each morning. In the middle of the afternoon over a period of 26 days she measured the following outcome variables: (1) her weight; (2) her score in Tetris; (3) her mood on a 0-5 scale; (4) her energy; and (5) whether she got a question right on the math GRE.

```
d <- read.dta("./data/Hough_WorkingPaper_2010.dta")
head(d)
```

```
##   day run weight tetris mood energy appetite gre
## 1   1   1     21  11092    3      3        0   1
## 2   2   1     21  14745    3      1        2   0
## 3   3   0     20  11558    3      3        0   1
## 4   4   0     21  11747    3      1        1   1
## 5   5   0     21  14319    2      3        3   1
## 6   6   1     19   7126    3      2        0   1
```

a. Suppose you were seeking to estimate the average effect of running on her Tetris score. Explain the assumptions needed to identify this causal effect based on this within-subjects design. Are these assumptions plausible in this case? What special concerns arise due to the fact that the subject was conducting the study, undergoing the treatments, and measuring her own outcomes?

b. Estimate the effect of running on Tetris score. Use regression to test the sharp null hypothesis that running has no immediate or lagged effect on Tetris scores. (**Note** the book calls for randomization inference, but this is a bit of a tough coding problem. **HINT**: For the second part of part (b), run one regression of today's score on both today's treatment assignment and yesterday's treatment assignment. Then, calculate the p-value that both effects are zero.)

c. One way to lend credibility to with-subjects results is to verify the no-anticipation assumption. Use the variable `run` to predict the `tetris` score *on the preeceding day*. Presume that the randomization is fixed. Why is this a test of the no-anticipation assumption? Does a test for no-anticipation confirm this assumption?

d. If Tetris responds to exercise, one might suppose that energy levels and GRE scores would as well. Are these hypotheses borne out by the data?

e. **Additional Mandatory Question**: Note that the observations in this regression are not necessarily all independent of each other. In particular, think about what happens when we lag a variable. Given this non-independence, would you expect randomization inference to give you a better answer than the regression answer you just obtained in (b)? Why? Which number(s) do you expect to be different in regression than in randomization inference? What is the direction of the bias? (This is a conceptual question, so you do not need to conduct the randomization inference to answer it. However, you are certainly welcome to try that exercise if you are curious.)