Problem Set 3

FILL IN YOUR NAME HERE

Question 1

Consider the following dataset

```
q1 <- tibble(
  group = c("Turnout Rate among those contacted",
            "Turnout Rate among those not contacted",
            "Overall Turnout Rate"),
  treatment = c(54.43, 36.48, 41.38),
  t_numbers = c(395, 1050, 1445),
  control = c(NA, 37.54, 37.54),
  c_numbers = c(NA, 5645, 5645)
)
```

Estimate the following quantities. Explain what each one means substantively.

$E[d_i(1)]$

$E[Y_i(0)|d_i(1) = 0]$

$E[Y_i(0)|d_i(1) = 1]$

$E[Y_i(1)|d_i(1) = 1]$

Question 2

Explain whether each of the following statements is true or false for the case of one-sided noncompliance. Assume that the experiment satisfies non-interference and excludability assumptions.

a) If the ITT is negative, the CACE must be negative.

b) The smaller the ITT_D, the larger the CACE

c) One cannot identify the CACE if no one in the experiment receives treatment.

Question 3

Consider the following paragraph describing an experiment:

"Researchers were interested the effect of encouraging college students at Berkeley to eat healthier. Canvassers visited dorms and encouraged students to eat more vegetables. Outcomes were measured by whether a student added at least one serving of vegetables to their next meal in the dining halls. When implementing the experiment, researchers encountered one-sided non-compliance. 1015 of the 1849 students assigned treatment were successfully contacted. None of the 1430 students assigned to control were contacted. 591 students in the treatment group added vegetables to their next meal, as opposed to 377 in the control group. 429 of the 1015 students who were successfully contacted add vegetables to their meal as opposed to 539 of the 2264 students that were not canvassed."

a) Estimate the ITT. What does the ITT mean in this context?

b) Estimate the ITT_D. What does it mean here?

c) Estimate the CACE and interpret the results.

d) Explain why comparing the vegetable eating rates of the treated and untreated subjects will produce misleading estimates of the ATE.

Question 4

Consider the general case of two-sided non-compliance. Assume that excludability and non-interference assumptions hold. State whether the following statements are true or false and why.

a) Among Compliers, the ITT equals the ATE.

b) Among Defiers, the ITT equals the ATE.

c) Among Always-Takers and Never-Takers, the ITT and ATE are zero.

Question 5

Suppose a sample contains 30% Always-Takers, 40% Never-Takers, 15% Compliers and 15% Defiers. What is the ITT_D?

Question 6

Hyde (2010) considers the effect of international election observers on monitoring election fraud. Due to difficulty in reaching villages and time constraints, observers monitored 68 of the 409 polling places assigned to treatment. Observers monitored 21 of the 1562 stations assigned to the control group. The primary outcome measure is the number of ballots declared invalid by polling station officials.

The dataset for this problem is called "Hyde_POP_2012.csv"

a) Is monotonicity a plausible assumption in this application?

b) Under the assumption of monotonicity, what proportion of polling locations are Compliers, Never-Takers, and Always-Takers?

c) Using the dataset, estimate the ITT and CACE. Interpret the results.

d) Use randomization to test the sharp null hypothesis that there is no intent-to-treat effect for any polling location. Interpret the results. Explain why testing this null hypothesis that the ITT is zero for all units is the same as testing the null hypothesis that the ATE is zero for all compliers.