# Assignment 2 Answer Key

Thomas Leavitt

14 August 2021

## Question 1

If we assume that the IV assumptions are valid for the Round 2 respondents in the Albertson and Lawrence (2009) study, we can estimate the average causal effect of encouragement, $\mathbf{Z}$, on dose, $\mathbf{D}$, via the estimator of the $ITT_D$, which is as follows:

$$\begin{aligned}\widehat{ITT_D} &= \frac{\mathbf{Z}'\mathbf{D}}{\mathbf{Z}'\mathbf{Z}} - \frac{(\mathbf{1}-\mathbf{Z})'\mathbf{D}}{(\mathbf{1}-\mathbf{Z})'(\mathbf{1}-\mathbf{Z})} \\ &= \left(\frac{1}{\sum_{i=1}^{n} Z_i}\right)\sum_{i=1}^{n} Z_i D_i - \left(\frac{1}{\sum_{i=1}^{n}(1-Z_i)}\right)\sum_{i=1}^{n}(1-Z_i)\,D_i\end{aligned}$$

For the first study, the estimate of the effect of encouragement on program viewing is:

$$\frac{244}{510} - \frac{74}{579} \approx 0.3506.$$

For the second study, the estimate is:

$$\frac{117}{259} - \frac{11}{248} \approx 0.4074.$$

## Question 2

A hypothetical outcome variable one could collect is "Views on Addiction Policy,'' measured on a scale from 0 – 10 in which 0 represents no support for public treatment programs and 10 is full support. Let's imagine that we have the following results:

$$\begin{aligned}\frac{3723}{510} - \frac{2895}{579} &= 7.3 - 5 \\ &= 2.3.\end{aligned}$$

The CACE estimator is $\frac{\widehat{\text{ITT}}}{\widehat{\text{ITT}_D}}$, where the $\widehat{\text{ITT}_D}$ is defined as above and the $\widehat{\text{ITT}}$ is defined as follows:

$$\begin{aligned}\widehat{ITT_D} &= \frac{\mathbf{Z}'\mathbf{Y}}{\mathbf{Z}'\mathbf{Z}} - \frac{(\mathbf{1}-\mathbf{Z})'\mathbf{Y}}{(\mathbf{1}-\mathbf{Z})'(\mathbf{1}-\mathbf{Z})} \\ &= \left(\frac{1}{\sum_{i=1}^{n} Z_i}\right)\sum_{i=1}^{n} Z_i Y_i - \left(\frac{1}{\sum_{i=1}^{n}(1-Z_i)}\right)\sum_{i=1}^{n}(1-Z_i)\,Y_i\end{aligned}$$

Using the hypothetical results described earlier, out CACE estimate is simply the ratio of our estimate of the ITT and the $ITT_D$, which is:

$$\frac{2.3}{0.3506} = 6.559718.$$

# Question 3

The Round 1 sample recruitment consisted of a survey conducted on a random sample of individuals from five metropolitan areas. If the researcher wants to make causal inferences among only the individuals who responded to the initial sample recruitment, then knowledge of the response rates are not necessary.

# Question 4

Albertson and Lawrence (2009) state that the study they analyzed used random sampling from five metropolitan areas to recruit individuals into the study. Among individuals who were successfully recruited, the researchers assigned units to treatment and control groups. Among the individuals in this study population, the researchers were able to contact only 80% of them to measure outcomes in Round 2 (Albertson and Lawrence, 2009, 284).

Since we know that there were 1089 individuals who responded in Round 2 and that these individuals make up 80% of the round 1 study popuation, we can reason backwards to infer that there were $\frac{1089}{0.8} \approx$ 1361 individuals in the study, only 1089 of which responded in Round 2.

Let $r_{Ti}$ be an indicator for whether subject $i$ would respond in Round 2 if assigned to treatment and let $r_{Ci}$ be an indicator for whether subject $i$ would respond in Round 2 if assigned to control.

The observed reporting status of unit $i$ is $R_i = Z_i r_{Ti} + (1 - Z_i) r_{Ci}$. If $R_i = 1$, then the researcher will observe $y_{Ci}$ for unit $i$ if $Z_i = 0$ and $y_{Ti}$ for unit $i$ if $Z_i = 1$. By contrast, if $R_i = 0$, then $Y_i$ will be unobserved.

Now we can define four distinct types of subjects with regard to attrition:

| $z_i = 0$ | $z_i = 1$ | Type of Subject |
|---|---|---|
| $r_{Ci} = 1$ | $r_{Ti} = 1$ | *Always-Reporter* (AR) |
| $r_{Ci} = 0$ | $r_{Ti} = 1$ | *If-Treated-Reporter* (ITR) |
| $r_{Ci} = 1$ | $r_{Ti} = 0$ | *If-Untreated-Reporter* (IUR) |
| $r_{Ci} = 0$ | $r_{Ti} = 0$ | *Never-Reporter* (NR) |

Let $s_i = s \in \{\text{AR}, \text{ITR}, \text{IUR}, \text{NR}\}$ denote the principal stratum of the $i$th unit and let $n^s$ denote the (typically unknown) number of units in each stratum. We analogously denote the average of treated and control potential outcomes in each stratum by $\bar{y}_T^s$ and $\bar{y}_C^s$.

Albertson and Lawrence (2009) measure $D_i$ and $Y_i$ among only 80% of the individuals who were assigned to treatment and control. The individuals who report their $D_i$ and $Y_i$ outcomes could be "Always-Reporters," "If-Treated-Reporters" or "If-Untreated-Reporters." Albertson and Lawrence (2009) restrict their analysis to only the individuals whose $D_i$ and $Y_i$ outcomes they did observe.

## Question 4a

Despite proper use of random assignment procedures during Round 1, the possibility that, for many individuals, $r_T = 1$ while $r_C = 0$ poses a threat to inference among Round 2 respondents. To see why, consider, e.g., the ITT estimator under attrition:

$$\begin{aligned}\hat{\tau} &= \hat{\tau}_T - \hat{\tau}_C \\ &= \frac{\sum_{i=1}^{N} Z_i R_i Y_i}{\sum_{i=1}^{N} Z_i R_i} - \frac{\sum_{i=1}^{N} (1 - Z_i) R_i Y_i}{\sum_{i=1}^{N} (1 - Z_i) R_i}\end{aligned}$$

To derive the expected value of $\mathrm{E}\left[\hat{\tau}\right]$, first consider only the term $\hat{\tau}_T$, which is a ratio of two random quantities. The denominator is the number of units with $r_{Ti} = 1$ that are randomized into the treatment group. The numerator is the sum of treated potential outcomes among units with $r_{Ti} = 1$ who are randomized into the treatment group. Even though $\sum_{i=1} Z_i$ is fixed across all assignments, $\sum_{i=1}^{N} Z_i R_i$ is not. Let $W_i = Z_i R_i = Z_i r_{Ti}$, where $W_i = 1$ if $Z_i = 1$ and $r_{Ti} = 1$ and $W_i = 0$ otherwise, and denote the number of units with $r_{Ti} = 1$ randomized into the treatment group by $W_T$. The quantity $W_T$ is a random variable that can take on the value of an integer $w_T \in \mathcal{W}_T = \left\{\max\left(0, n^{\mathrm{AR}} + n^{\mathrm{ITR}} - n_C\right), \ldots, \min\left(n_T, n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)\right\}$.

Then we can appeal to the law of iterated expectation to write $\mathrm{E}\left[\hat{\tau}_T\right]$ as

$$\mathrm{E}\left[\hat{\tau}_T\right] = \sum_{w_T \in \mathcal{W}_T} \mathrm{E}\left[\hat{\tau}_T \mid W_T = w_T\right] \Pr\left(W_T = w_T\right).$$

It is easy to show that $\mathrm{E}\left[\hat{\tau}_T \mid W_T = w_T\right] = \dfrac{1}{n^{\mathrm{AR}} + n^{\mathrm{ITR}}} \sum_{i: r_{Ti} = 1} y_{Ti}$ for any value of $w_T \in \mathcal{W}_T$. Hence, by the law of total probability,

$$\begin{aligned}\mathrm{E}\left[\hat{\tau}_T\right] &= \sum_{w_T \in \mathcal{W}_T} \mathrm{E}\left[\hat{\tau}_T \mid W_T = w_T\right] \Pr\left(W_T = w_T\right) \\ &= \mathrm{E}\left[\hat{\tau}_T \mid W_T = w_T\right] \sum_{w_T \in \mathcal{W}_T} \Pr\left(W_T = w_T\right) \\ &= \mathrm{E}\left[\hat{\tau}_T \mid W_T = w_T\right] (1) \\ &= \frac{1}{\left(n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)} \sum_{i: r_{Ti} = 1} y_{Ti} \\ &= \frac{1}{\left(n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)} \left( \sum_{i: s_i = \mathrm{AR}} y_{Ti} + \sum_{i: s_i = \mathrm{ITR}} y_{Ti} \right) \\ &= \frac{1}{\left(n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)} \sum_{i: s_i = \mathrm{AR}} y_{Ti} + \frac{1}{\left(n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)} \sum_{i: s_i = \mathrm{ITR}} y_{Ti} \\ &= \frac{n^{\mathrm{AR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)} \frac{1}{n^{\mathrm{AR}}} \sum_{i: s_i = \mathrm{AR}} y_{Ti} + \frac{n^{\mathrm{ITR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)} \frac{1}{n^{\mathrm{ITR}}} \sum_{i: s_i = \mathrm{ITR}} y_{Ti} \\ &= \frac{n^{\mathrm{AR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)} \bar{y}_T^{\mathrm{AR}} + \frac{n^{\mathrm{ITR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)} \bar{y}_T^{\mathrm{ITR}}.\end{aligned}$$

Analogous logic for $\hat{\tau}_C$ implies that

$$\begin{aligned}\mathrm{E}\left[\hat{\tau}_C\right] &= \frac{1}{\left(n^{\mathrm{AR}} + n^{\mathrm{IUR}}\right)} \sum_{i:r_{Ci}=1} y_{Ci} \\ &= \frac{n^{\mathrm{AR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{IUR}}\right)} \bar{y}_C^{\mathrm{AR}} + \frac{n^{\mathrm{IUR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{IUR}}\right)} \bar{y}_C^{\mathrm{IUR}}.\end{aligned}$$

Hence, the expected value of the ITT estimator, $\hat{\tau}$, is

$$\mathrm{E}\left[\hat{\tau}\right] = \frac{n^{\mathrm{AR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)} \bar{y}_T^{\mathrm{AR}} + \frac{n^{\mathrm{ITR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)} \bar{y}_T^{\mathrm{ITR}} - \frac{n^{\mathrm{AR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{IUR}}\right)} \bar{y}_C^{\mathrm{AR}} + \frac{n^{\mathrm{IUR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{IUR}}\right)} \bar{y}_C^{\mathrm{IUR}}. \tag{1}$$

Under random assignment, the first term, $\hat{\tau}_T$, is unbiased for the average of treated potential outcomes among Always-Reporters and If-Treated-Reporters, and the second term, $\hat{\tau}_C$, is unbiased for the average of control potential outcomes among Always-Reporters and If-Untreated-Reporters. The expected value of the ITT estimator is no longer equal to a meaningful average causal effect; rather the expected value of the estimator is equal to the difference in average treated potential outcomes among one mixture of units minus the average control potential outcome among another mixture of units.

The same logic also applies to the $\mathrm{ITT}_D$ in which treatment compliance, $D_i$, is the relevant outcome of units $i$ instead of $Y_i$.

## Question 4b

Referring back to (1), if $r_{Ti} = r_{Ci}$ for all Round 2 units, then the proportions of If-Treated-Reporters and If-Untreated_reporters are 0. Hence, the expected value of the Difference-in-Means estimator is

$$\begin{aligned}\mathrm{E}\left[\hat{\tau}\right] &= \frac{n^{\mathrm{AR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)} \bar{y}_T^{\mathrm{AR}} + \frac{n^{\mathrm{ITR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{ITR}}\right)} \bar{y}_T^{\mathrm{ITR}} - \frac{n^{\mathrm{AR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{IUR}}\right)} \bar{y}_C^{\mathrm{AR}} + \frac{n^{\mathrm{IUR}}}{\left(n^{\mathrm{AR}} + n^{\mathrm{IUR}}\right)} \bar{y}_C^{\mathrm{IUR}} \\ &= \bar{y}_T^{\mathrm{AR}} - \bar{y}_C^{\mathrm{AR}}\end{aligned}$$

which now is a well-defined causal quantity — the average causal effect among Always-Reporters.

As in Question 4a, the same logic also applies to the $\widehat{\mathrm{ITT}}_D$ in which treatment compliance, $D_i$, is the relevant outcome of units $i$ instead of $Y_i$. So long as $r_{Ci} = r_{Ti}$, if the researcher simply omits the units whose outcomes were not observed, as Albertson and Lawrence (2009) do, then inference remains valid among the subpopulation of "Always-Reporters."

# Question 5

The strength of an instrument is equivalent to the proportion of Compliers. Whether a unit is a Complier or not depends on that unit's unobservable potential outcomes. Hence, we cannot actually know the proportion of Compliers in each of the respective studies. We can, however, unbiasedly estimate the proportion of Compliers under the assumption of No Defiers. In particular, the $\widehat{\mathrm{ITT}}_D$ estimator in Question 1 is unbiased for the proportion of Compliers in Study 1 and Study 2 of Albertson and Lawrence (2009). These two estimates are given in Question 1 and are approximately 0.35 and 0.41 for studies 1 and 2, respectively. Hence, Study 2 appears to have a stronger instrument.

Now let's assess the strength of the instrument in the Acorn study by Arceneaux (2005). To do so, we calculate the number of treated and control units on the individual (not precinct) level. We also calculate the number of individuals contacted in treatment and control groups. Then to estimate the proportion of Compliers, we take the difference in average number of people contacted in the treatment and control groups.

```
acorn_data <- read.csv("acorn03.csv")

(1/sum(acorn_data$z * acorn_data$size)) * sum(acorn_data$z * (acorn_data$size * acorn_data$contact))
```

```
[1] 0.6274073
```

```
(1/sum((1 - acorn_data$z) * acorn_data$size)) * sum((1 - acorn_data$z) * (acorn_data$size *
    acorn_data$contact))
```

```
[1] 0
```

The estimated effect of assignment to treatment (GOTV campaign) on receiving the treatment is approximately 0.63. Therefore, the Acorn study appears to have the strongest instrument. Also, notice that in the Acorn study, the assumption of No Defiers is true by design, which further strengthens the credibility of interpreting $\widehat{\text{ITT}}_D$ as unbiased estimator for the proportion of Compliers.

# References

Albertson, Bethany and Adria Lawrence. 2009. "After the Credits Roll: The Long-Term Effects of Educational Television on Public Knowledge and Attitudes." *American Politics Research* 37(2):275–300.

Arceneaux, Kevin. 2005. "Using Cluster Randomized Field Experiments to Study Voting Behavior." *The Annals of the American Academy of Political and Social Science* 601(1):169–179.