Causal Inference Final Paper

Post-Treatment Bias

“Conditioning on post-treatment variables eliminates the advantages of randomization because we are now comparing dissimilar groups.” (5)

“concerns about post-treatment bias are not really (or only) about the post-treatment variable itself. The problem is that by conditioning on a post-treatment variable we have unbalanced the treatment and control groups with respect to every other possible confounder.” (6)

“In short, when we include a post-treatment variable in the set of conditioning variables either directly or indirectly, Assumption (1) is violated. As a result, ⌧ 6= for the reasons discussed above. Standard estimates such as the di↵erence in conditional means (ˆ ) will therefore be biased regardless of sample size, measurement precision, or estimation method.9 Further, the bias of standard estimates such as ˆ can be in any direction and of any magnitude depending on the value of unknown (and unknowable) parameters (e.g., Y , the e↵ect of the unmeasured confounder on the outcome). Once we have conditioned on a post-treatment variable, we have eliminated the assurance of unconfoundedness provided by randomization.” (11)

“the post-treatment covariate (x) and the outcome (y) share an unmeasured cause (u)”

“In our example, we might wish to estimate the e↵ect of the civics education class only among low-interest students to show that the e↵ect is not isolated to previously engaged students. Dropping respondents based on manipulation checks is often done to show that the estimated treatment e↵ect is larger among compliers, which might appear to suggest that the treatment is working through the researchers’ proposed mechanism. This reasoning is wrong. Selecting a portion of the data based on post-treatment criteria will not allow us to generate an unbiased estimate of the treatment e↵ect within an interesting subset of respondents. Instead, we will obtain a biased estimate among an endogenously selected group.” (13)

“By selecting based on a criterion that is partially a function of unobserved covariates and the treatment, we have inadvertently created imbalance in the treatment and control conditions with respect to u.” (14)

“Including control variables is therefore potentially appropriate, but only covariates that are unrelated to the treatment and preferably measured in advance” (23)

“Moderators that are vulnerable to treatment spillovers like racial resentment should be measured pre-treatment” (23)

Key Assumptions- IV

1. Ignorability
   1. assume Z, our instrument is randomly assigned
2. Exclusion Restriction
   1. exclusion restriction says that if your treatment wouldn’t be different even if your instrument assignment was different, then your outcome (or, more generally, the distribution of your outcome) also won’t be different
   2. possible that encouragement affected grades even if students didn’t attend lab—could have been more likely to attend office hours, increased amount of study time
   3. A colloquial way of phrasing this is that the instrument only affects the outcome through the treatment
3. Monotonicity
   1. No defiers
4. Non-zero correlation between instrument and treatment
   1. An instrument isn’t useful if it doesn’t actually predict the treatment
5. SUTVA
   1. one person’s treatment assignment would not affect another’s outcome

Key Assumptions- Propensity Scores

1. Ignorability of the treatment assignment
2. Balance
3. Overlap
4. SUTVA

Ways Post-Treatment Bias is introduced

* Condition on a variable that is affected by the treatment
  + i.e., measure interest after participation in a program that may affect interest
  + effort to prevent omitted variable bias
  + account for noncompliance
  + measure a moderator after experimental manipulation
* Dropping or selecting observations based on criteria influenced by the treatment
  + Sometimes unavoidable
  + “The treatment itself may cause some respondents to be more likely to be omitted from the sample, a phenomenon which is usually termed non-random attrition.” (12)
  + “researchers frequently drop subjects who fail a post-treatment manipulation check or other measure of attention or compliance” (13)
    - “can imbalance the sample with respect to observed or unobserved confounders” (13)

“Simulating the Effect of Controlling for Post-Treatment Variables on Treatment Estimates”

Shannon Kay

APSTA 2012 Causal Inference

**Motivation**

While it seems obvious that including variables affected by an experimental treatment when analyzing results might bias the estimate of a treatment, the reality is not as straightforward. One of the main motivations for researching the effects of post-treatment variables comes from my experience working as a Research Assistant at the Marron Institute. Oftentimes, our practitioners are interested in what they consider the true effect of the treatment—the difference in outcomes for those in the treatment group as adjusted for dosage or another mediating factor, rather than simply looking at an intent to treat analysis of those who were assigned to the treatment condition. In some trials, there may be a valid method to complete this analysis with appropriate statistical rigor. More often, though, unless this possibility is considered during the trial’s design, intent to treat is the only possible method to make any causal attributions. The problem, however, arises when we are asked to subset the data based on a post-treatment variable, or include a non-randomized dosage measure in the analysis.

When treatment groups are randomized, we assume ignorability—that once randomized, the treatment and control groups are equal on all measured and unmeasured pre-treatment characteristics. The addition of post-treatment variables breaks the ignorability assumption by controlling for a variable that was not taken into consideration during randomization. To understand how the inclusion of post-treatment variables influences treatment effect estimates, I imagine a hypothetical situation where 1000 NYU students enrolled in Stats 101 are randomly encouraged to attend an optional lab to supplement their statistics class. Random assignment generates treatment and control groups that are evenly distributed with regard to gender and age. Students who are not encouraged may choose to attend the lab but must find the information independently. There are 10 lab sessions. Students’ final grades for Stats 101 are recorded at the end of the semester, as well as the number of lab sessions they attended. Attendance thus becomes our post-treatment variable.

To understand how controlling for attendance affects the estimate of the treatment effect, I will use an instrumental variable (IV) approach to estimate the Complier Average Causal Effect (CACE) and two propensity score methods to estimate the Average Treatment Effect on the Treated (ATT). Since the instrumental variable method requires an instrument as well as treatment, which the propensity score estimators require only treatment status, I will simulate data using two data-generating processes. All methods will compare three post-treatment attendance variables of different strength correlation to the instrument for instrumental variables, and to treatment for propensity scores. I choose to investigate these estimands because they are most comparable to the interests of our practitioners.

**Assumptions**

*Instrumental Variables*

The first assumption behind instrumental variables is *ignorability*, which means that the instrument is randomly assigned. Random assignment allows us to assume that the groups generated are equal on all observed and unobserved confounders. With this established, we can reasonably attribute any differences in outcomes to the treatment. In the case of instrumental variables, assignment to encouragement, the instrument, is randomized.

In instrumental variables, there are four possible subject types. The never-takers will not participate in the treatment regardless of whether they are offered the treatment. The compliers will participate if they are offered the treatment, and will not participate if they are not offered the treatment. Always-takers will participate in the treatment regardless of being offered the treatment (i.e., they will seek out and obtain the treatment on their own). Defiers will do the opposite of their treatment assignment; if they are offered the treatment, they will refuse it, but if they are not offered the treatment, they will seek it out and obtain it. Our second assumption, *monotonicity*, assumes that there are no defiers.

From there, we assume *exclusion restriction*. Within our defined subject types, always-takers and never-takers will always be treated or untreated, regardless of instrument assignment (encouragement)—the always takers will always find a way to receive the treatment, and never-takers will never accept treatment. By exclusion treatment, if your treatment condition remains the same under either instrument assignment, you will have the same outcome under either instrument assignment. This implies that always-takers and never-takers will have the same final grade in Stats 101 whether or not they are encouraged to attend the optional lab sessions.

An important assumption for instrumental variables is a *non-zero correlation between the instrument and the treatment*. In this case, encouragement must correlate

Lastly, we assume *SUTVA,* or the *stable unit treatment value assumption*. Colloquially, the treatment assignment of one subject will not affect the outcome of another subject. In terms of encouragement to go to an optional, supplementary lab, this means assuming that those who are encouraged to go do not change the treatment assignment of another student by bringing friends who were not encouraged to go to the lab sessions, and that …….

*Propensity Scores*

Propensity score models also assume ignorability of the treatment assignment (rather than instrument assignment) and SUTVA

1. Ignorability of the treatment assignment
2. Balance
3. Overlap
4. SUTVA

**Designs & Estimators**

*Estimator 1: Two-Stage Least Squares Regression*

In Two-Stage Least Squares (TSLS), we estimate the CACE by first regressing whether or not students attended supplementary lab sessions (treatment) on whether or not they were encouraged to attend the labs (instrument). The coefficient for on the instrument provides an estimate of the percentage of compliers in our sample. We then regress final Stats 101 grades on the treatment along with any other relevant predictors. The resulting coefficient on treatment is the estimated Complier Average Causal Effect. While TSLS is known to provide incorrect estimates of standard errors because it does not account for correlation between the two regression equations

*Estimator 2: Propensity Score Weighting using MatchIt*

MatchIt uses

*Estimator 3: Propensity Score Weighting using PS (Twang Package)*

**Simulation Set-Up**

In this simulation, I aim to explore treatment estimation methods’ reliance on the ignorability assumption by controlling for post-treatment variables in my models.

For simplicity, the covariates included in the model are already balanced across groups; this allows me to isolate the ignorability assumption via the inclusion of the post-treatment variable, as

I generate three post-treatment attendance variables that are correlated with compliance in the case of instrumental variables, or with treatment in the case of propensity score matching. Attendance is a proportion of the possible number of lab sessions. If the attendance variable is 1, the student attended all 10 sessions, if attendance is 0, the student did not attend any lab sessions.

I vary the degree of correlation by using a random beta distribution where alpha is held at 2 while beta is manipulated.

For instrumental variables, all three attend variables are always 1, or 100% attendance, for the always-takers and always 0, or never attended. Compliers who are not randomly assigned encouragement also always receive 0’s for attendance. The manipulation is contained to compliers who are randomly assigned to encouragement. When the

In Attend 1, beta is equal to 15; this assigns the majority of the compliers an attendance rate under .3, or 3 labs, and potentially makes the compliers look more like the never-takers.

In Attend 2, beta is equal to 2, which

In Attend 3, beta is equal to .5, which places the majority of compliers at an attendance rate above

For propensity scores, manipulation of post-treatment variables is relegated to those randomly assigned to the treatment. All non-treated subjects receive a 0 for attendance. The distribution for the three attendance variables is manipulated in the same manner as in the IV data generating progress.

**Simulation Results**

**Discussion**