Assignment #4

ECON 5783 — University of Arkansas

Prof. Kyle Butts

These assignments should be completed in groups of 1 or 2 but submitted individually. My preference is for you to use Rmarkdown files to have your code, results, and your answers to the questions intermixed. Since I am not requiring you to code in R for these assignments, you can use latex or microsoft word to write up answers alternatively. Not that in these cases, I would like you to upload your code seperately.

Theoretical Questions

1. Say you believe that conditional on age, treatment is randomly assigned $(Y_i(1), Y_i(0)) \perp\!\!\!\perp D_i \mid \text{Age}_i$.

    i. How would you estimate the conditional average treatment effect for 30 year olds?

    ii. Say you repeat this for all the ages in your sample (say 25-55 years old). How would you aggregate these to the overall average treatment effect?

2. For the following 3 examples, tell me if you believe the conditional independence assumption. If you do not, try to 'tell me a story' of selection bias (i.e. give an example of some unobservable confounder).

    i. A researcher is trying to estimate the returns to taking an online coding class. Since they are concerned about selection into treatment, they condition on having a college degree (a proxy for 'motivation').

    ii. A researcher wants to estimate the effect of new apartment construction on home prices. They think that conditional on a neighborhood's average income, which neighborhoods get a new apartment is randomly assigned.

    iii. A researcher wants to estimate the impact of smoking on later life health outcomes. They are concerned smoking is correlated with other risky behavior, so you want to match individuals based on their average amount of weekly drinking.

Coding Exercise

We are going to try using some conditional independence estimators today. We will use the Almond, Chay, Lee (2005, QJE) dataset thinking about mother's smoking on child's birthweight (you can use `haven::read_dta` to open `data/almond_chay_lee.dta`.). The variables are listed below. Our outcome is `bweight` and treatment is `mbsmoke`. A complete codebook is at the end of this document

1. Let's try to make an argument that the unconditional independence assumption is not valid. To do so, let's conduct a balance check. You can follow the steps of the 'Check Initial Imbalance' section of https://kosukeimai.github.io/MatchIt/articles/MatchIt.html#check-initial-imbalance conduct a balance check using some important covariates (pick a few) to see if treatment is correlated with these variables. Report your findings.

2. Now let's try to perform a matching estimator. To do so, you will use the `MatchIt` package. Use the `matchit()` function to perform nearest neighbor matching. Let's match exactly on `mmaried`, `mhisp`, and `alcohol`. Let's nearest neighbor match for `mage` and `medu`.
   Then, get the matched dataset using `match.data()` on the result of `matchit()`. With the matched dataset, estimate the difference-in-means estimator. Interpret your estimate in words

3. Now let's estimate a regression adjustment by hand. Use the same varaibles as before: `mmaried`, `mhisp`, `alcohol`, `mage`, and `medu`.

4. Last, we will estimate the IPTW esitmator using the `WeightIt` package. Use `weightit()` to estimate propensity scores.

   i. Check for balance after weighting by the estimate propensity score using the `bal.tab()` function. Comment on the results.

   ii. Using `lm_weightit()` function to estimate the IPW estimator using the results from `weightit()`. See the first example in the help menu of `lm_weightit()` for details.

Table 1 — Almond, Chay, and Lee (2005) Dataset

Variable	Description
bweight	infant birthweight (grams)
mmarried	1 if mother married
mhisp	1 if mother hispanic
fhisp	1 if father hispanic
foreign	1 if mother born abroad
alcohol	1 if alcohol consumed during pregnancy
deadkids	previous births where newborn died
mage	mother's age
medu	mother's education attainment
fage	father's age
fedu	father's education attainment
nprenatal	number of prenatal care visits
monthslb	months since last birth
order	order of birth of the infant
msmoke	cigarettes smoked during pregnancy
mbsmoke	1 if mother smoked
mrace	1 if mother is white
frace	1 if father is white
prenatal	trimester of first prenatal care visit
birthmonth	month of birth
lbweight	1 if low birthweight baby
fbaby	1 if first baby
prenatal1	1 if first prenatal visit in 1 trimester