MaCSS222, Spring 2025

Professor Bryan Graham

Problem Set 3

Due: On date agreed to in class on 3/10

Problem sets are due at 11:59PM of the due date. The reader will provide instructions on how to turn in your problem set. You may work in groups, but each student should turn in their own write-up (including a "printout" of a narrated/commented and executed Jupyter Notebook if applicable). Please include a list of classmates you collaborated with when you turn in your problem set. Please also e-mail a copy of any Jupyter Notebook to the GSI (if applicable).

1 Covariate adjustment via the propensity score: build-up

In this problem set we will analyze the replication dataset for McEwan et al. (2015) (this Stata format dataset can be found on GitHub and is called **sat_eepa_rep.dta**). Please read this article first; the context will be helpful as you complete this assignment. McEwan et al. (2015) study the efficacy of the alternative rural secondary education program Sistema de Aprendizaje Tutorial (SAT) relative to Centros de Educacion Basicos (CEB). The population of interest is rural youth in five departments of Honduras who have completed 6th grade (and are thus eligible to enroll in secondary school). You can find a Python Jupyter notebook stub on GitHub that should help you get started on the problem set.

Let $X_i = 1$ if (randomly sampled) youth i lives in a village where the SAT program is available and $X_i = 0$ if, instead, they live in a village with a Centro de Educacion Basicos (CEB). Let $Y_i(1)$ be the potential outcome when, possibly contrary to fact, youth i has access to SAT and $Y_i(0)$ their potential outcome if instead (again possibly contrary to fact) they have access to a CEB.

The researcher observes

$$Y_i = (1 - X_i) Y_i(0) + X_i Y_i(1), \tag{1}$$

since it is logically impossible to observe both $Y_i(1)$ and $Y_i(0)$ for a single unit at the same time. The McEwan et al. (2015) replication files include data from 94 villages, 47 with CEBs and 47 with SAT groups. In each village the entire cohort on graduating 6th graders in 2008 was interviewed for a total sample size of 1,426 youth. For concreteness consider the outcome of interest, Y_i, to be measured mathematics knowledge at the end of 2010. This would coincide with the end of 8th grade for those students choosing to enroll in secondary school. The W_i denote a vector of baseline variables measured for each respondent in 2008.

1. Imagine – although this was not the case – that the 47 villages with SAT programs were chosen at random with the remaining villages given CEBs. Explain in *words* why random assignment ensures that

$$Y_i(0), Y_i(1) \perp X_i. \tag{2}$$

Explain in words the following two implications of (2)

$$\begin{aligned} \mathbb{E}[Y_i(1) | X_i = 1] &= \mathbb{E}[Y_i(1) | X_i = 0] = \mathbb{E}[Y_i(1)] \\ \mathbb{E}[Y_i(0) | X_i = 1] &= \mathbb{E}[Y_i(0) | X_i = 0] = \mathbb{E}[Y_i(0)] \end{aligned} \tag{3}$$

Write down what both lines in equation (3) are saying in words. Why does treatment X_i have no predictive power for both $Y_i(1)$ and $Y_i(0)$ in the randomized control trial (RCT) setting?

2. Use (1) and (3) to show that

$$\mathbb{E}[Y|X_i = 1] = \mathbb{E}[Y_i(1)]$$

as well as that

$$\mathbb{E}[Y|X_i = 0] = \mathbb{E}[Y_i(0)],$$

and hence that

$$\begin{aligned}\beta^{\text{ATE}} &= \mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)] \\ &= \mathbb{E}[Y|X_i = 1] - \mathbb{E}[Y|X_i = 0]\end{aligned}$$

3. In practice neither SATs nor CEBs were randomly assigned to villages. Give 3 concrete reasons why, in the context studied by McEwan et al. (2015), it may be the case that

$$\begin{aligned}\mathbb{E}[Y_i(1)|X_i = 1] &\neq \mathbb{E}[Y_i(1)|X_i = 0] \\ \mathbb{E}[Y_i(0)|X_i = 1] &\neq \mathbb{E}[Y_i(0)|X_i = 0]\end{aligned}$$

Hint: Think about how the Honduran Ministry of Education might decide where to place CEBs. Are "CEB villages", and the youths residing in them, systematically different from non-CEB villages? How might these differences be related to the two potential outcomes, $Y_i(1)$ and $Y_i(0)$? Similarly, how does the La Asociación Bayán choose where to start SAT groups? Are "SAT villages", and the youths residing in them, systematically different from non-SAT villages? How might these differences be related to the two potential outcomes, $Y_i(1)$ and $Y_i(0)$?

4. McEwan et al. (2015) argue that, although (2) may not hold in the setting they study, the weaker restriction

$$Y_i(0), Y_i(1) \perp X_i|W_i \tag{4}$$

does hold. Describe (4) in plain words and assess its plausibility in the setting studied by McEwan et al. (2015). Provide both a positive and negative case, as well as your own bottom-line assessment.

2 Covariate adjustment via the propensity score: application

1. McEwan et al. (2015) base their analysis on the following implication of (4) above:

$$Y_i(0), Y_i(1) \perp X_i|e_0(W_i) \tag{5}$$

where $e_0(W_i) = \Pr(X_i = 1|W_i = w)$ is the propensity score. Give a narrative definition of the propensity score. Describe (5) in plain words.

2. Reproduce Table 2 of McEwan et al. (2015). What facts are the authors attempting to establish in this table? Why are they important for their analysis (Hint: consider your answers to Questions 2 and 3 in Part 1 above). If you could include additional variables in this table what would they be and why?

3. Reproduces Panels D and E of Table 3? Describe in plain words how the analysis in this table leverages (4) and (5) in different ways.

4. How do you think the McEwan et al. (2015) could be improved (e.g., additional data that might have been collected).

3 Quick reactions

Read the paper by Gunter & Daly (2012) and provide a few paragraphs of reactions. Do you think the evidence present in this paper is actionable? How could the study have been improved?

4 Further reading

There is a massive literature on propensity score analysis. This is good and bad. Good in that much material is available, bad in that curation of good resources can be difficult. The papers by Austin (2011) and Shrank et al. (2011) are non-technical and oriented towards health care professionals. You might find the latter especially helpful for Part 1 of this problem set. An elegant, but advanced, treatment is provided byImbens & Rubin (2015).

References

Austin, P. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3), 399 – 424.

Gunter, W. D. & Daly, K. (2012). Causal or spurious: using propensity score matching to detangle the relationship between violent video games and violent behavior. *Computers in Human Behavior*, 28(4), 1348 – 1355.

Imbens, G. W. & Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* Cambridge: Cambridge University Press.

McEwan, P. J., Murphy-Graham, E., Irribarra, D. T., Aguilar, C., & Rápalo, R. (2015). Improving middle school quality in poor countries: evidence from the honduras sistema de aprendizaje tutorial. *Educational Evaluation and Policy Analysis*, 37(1), 113 – 137.

Shrank, W. H., Patrick, A. R., & Brookhart, M. A. (2011). Healthy user and related biases in observational studies of preventive interventions: a primer for physicians. *Journal of General Internal Medicine*, 26(5), 546 – 550.