Q&A Chat Webinar 7 (Propensity Score Matching)

May 23rd, 2024 (Theory)

#	Question	Answer
1	Can you add RCT method to the Quantitative analysis?	Hi Desalegn, we'll pass this onto the presenter for his consideration. Thanks!
2	Please include Regression Discontinuity in this webinar series.	Hi, we'll pass this onto the presenter for his consideration. Thanks!
3	Our program participants are selected from localized villages, then matching with national data is appropriate and can the result be generalized to national populational	Good point. But the question on generalizability (or what is called external validity) is a different issue from matching.
4	Must the two groups have the same number of observations?	Not necessarily. We will learn latter about the idea of "common support". Essentially, you would usually want a larger sample of comparison group to choose matching observations from.
5	"I am working on my PhD Thesis; wellbeing inequality among internally displaced persons. Looking at the issue of self-selection with IDPs decision to either self-select into staying in camp or out of camp. I intend to use PSM to account for this type of endogeneity.	Try to follow the lecture today and see if you have additional questions. It seems like you have self-selection based on unobservable, so not the best scenario for PSM. Of course, PSM is better than nothing, still, you should understand the limitations of your design.
6	Please, can you explain again the selection assumption? I don't get it very well.	Propensity score matching assumes that once we control for all observable characteristics (i.e., conditional on X) then the two groups in our study are as good as random. This is a BIG assumption!!!
7	In some cases, journals require the use of sophisticated methods.	Sophisticated enough to accurately measure the missing counterfactual. Ultimately, if the assumptions do not hold, no amount of sophisticated econometrics will bail you out.
8	But for the propensity score method it is so hard to find exact match and we meet with intention to treatment for both group (treat & control; where some are included in treated but prefer not to be treated vice versa)	Yes. Exact matching is not easy to achieve.
9	Can there be insufficient matches of the scores??	correct me if this is not what you are asking. By insufficient matches, I assume you mean not finding a match for each treatment unit. We will discuss this scenario when we cover common support idea.
10	Is p hat the same as the predicted probability of logit or probit regression?	yes
11	How can we solve issues where the matches are not available or there are more matches for one score over the other (with treated and untreated)?	If you are not finding good matches, it implies your groups are too different from each other. Or your probit/logit regression is not good at predicting treatment assignment. The answer to both is better

		data and propensity score matching might not be suited in your situation
12	Yes Paul. Where not finding a match for each treatment... or a treatment score having more matches where others are not...	That's an issue of common support. Depending on your needs, you could choose to keep observations that are only within the region of common support (this would imply deleting observations without matches). If you do so, depending on your sample size, you might run into low power issues.
13	Most times my fear of overspecification makes me hesitate to throw in many variables. Any advice?	Do you best to predict treatment based on your understanding of how treatment is assigned/defined. We do no worry too much about overfitting because we are not looking at outcomes in our probit regression. The bigger issue with PSM is the UNOBSERVABLES you can't match on
14	What is the preferred solution based on the literature and your experience...?	So you need to obtain good matches. Start by adding variables linearly. Test your matches. If observations are too different between treated and matched untreated, then include quadratics. Match again. Test your matches. if still too different, add higher order terms and interactions. Match. Test your matches. Repeat
15	Can you share with us one paper for using the log of PS?	We are sharing the paper that is used for the example in the lecture.
16	Can we take log of Ps, on STATA by generating it?	yes, we will cover this next week.
17	Please, I hope Paul will share the code for making the graph.	tw hist Y if X==1, freq lcolor(sand) || hist Y if X==0, freq fcolor(none) lcolor(black)
18	In this case where the control group is larger than the treatment group can we apply the inverse probability weighting method	yes, you can always apply IPW method. But you still need to find good matches in the first place
19	Is PSM used in both experimental and non-experimental studies?	PSM is a non-experimental technique. In an RCT you would not really need to use matching.
20	I want to communicate further	Please send an email to airp3africa@air.org or jngao@air.org
21	What is the implication of using Log(OR) instead of the propensity scores?	In a famous article by Heckman, Ichimura and Todd in RESTUD 1997, the authors argue that using the log of the odds ratio is preferred when you have over-sampled from the T group. This is the case that we have, we have a dedicated sample of treated, much larger than their proportion in the population as a whole. If you are using national survey data for both T and comparison, then the proportions would be the same and you use the ps itself.
22	Please include RDD in the webinar series.	WE will likely have a dedicated series on impact evaluation and research design next year.