Problem Set 2

Background reading: Read the following articles in the order listed:

- Rosenbaum and Rubin (1983). Read slowly. You are expected to understand what the theorems are about.
- Blinder (1973) and Oaxaca (1973). Skim to make sure you know what the "Blinder-Oaxaca approach" is.
- Imbens (2004). Skim, and think of as a reference for the future. Then, after you have read lots more, come back and give it a more careful reading. Some good stuff there.
- DiNardo, Fortin and Lemieux (1996). Skim to get a flavor for the paper, but ignore the "pasting" parts unless you are excited about unions.
- Smith and Todd (2005) and Dehejia and Wahba (1999) and the other papers in the exchange (I think they are all on Rajeev Dehejia's website—he is at Tufts currently). These should be read quickly, but it is worth reflecting on the following question as you read: What exactly are these guys fighting about? Is it trivial, or is it a big deal?
- Busso, DiNardo and McCrary (2009). Some of the stuff I lectured on is drawn from an updated version of this manuscript, so if you got lost in lecture, this might be a handy reference.

After that, you should be ready to tackle the problem set. There is a lot to absorb, so you may still feel a bit muddled, but don't worry: the feeling fades with time and exposure.

The problem set is both computational and conceptual. Throughout, I assume you are using Stata, but you don't have to. If you are doing everything in some other package, that is fine. Adapt the instructions accordingly.

Question A

Suppose you have a dataset with 6 observations, with 3 treated ($T_i = 1$) and 3 control ($T_i = 0$) units. There is a single covariate, X_i, and an outcome Y_i. The 3 treated units have covariate values of 1, 2, and 3, and the 3 control units have covariate values of 4, 5, and 6. Suppose the model for the counterfactual outcomes is

$$Y_i(0) = \alpha_0 + \beta_0 X_i + \varepsilon_i^0 \quad (1)$$

$$Y_i(1) = \alpha_1 + \beta_1 X_i + \varepsilon_i^1 \quad (2)$$

and that while the pair $(Y_i(0), Y_i(1))$ is never observed, we observe the outcome

$$Y_i = T_i Y_i(1) + (1 - T_i) Y_i(0) \quad (3)$$

1. Is it possible to reweight the control observations so that they are similar to the treatment observations? Why or why not?
2. Load the 6 observations described into a Stata data set. Try to estimate a logit model in which you use X_i to predict the probability that $T_i = 1$. What happens? Discuss why this happens and what it means. Is the model identified?
3. (Note: Some parts are trick questions.) Suppose you want to estimate $\mathbb{E}[Y_i(1) - Y_i(0)|T_i = 1]$, or TOT. (a) Compute TOT assuming that for every value $x = 1, 2, \ldots, 6$, $P(X_i = x) = 1/6$. Set $\alpha_0 = \alpha_1 = 0$, $\beta_0 = 0.5$, and $\beta_1 = 1$ and further set $\mathbb{V}[\varepsilon_i^0] = \mathbb{V}[\varepsilon_i^1] = 0$, i.e., assume that $\varepsilon_i^0 = \varepsilon_i^1 = 0$ for each i. (b) What is the pair matching estimate for this case? Is this equal to the estimand? Why or why not? (d) What is the Blinder-Oaxaca estimate of TOT for this case? Is this equal to the estimand? Why or why not? (e) What is the reweighting estimate of TOT for this case? Is this equal to the estimand? Why or why not?

Question B

1. Write a Stata program that computes pair matching for TOT given inputs of Y_i, T_i, and X_i. If you don't know what this means, look over
`http://www.ssc.upenn.edu/scg/stata/stata-programming-1.ppt`
`http://www.stanford.edu/~roymill/cgi-bin/methods2010/material/stataCamp_1_ho.pdf`
`http://www.stanford.edu/~roymill/cgi-bin/methods2010/material/stataCamp_2_ho.pdf`
`http://www.stanford.edu/~roymill/cgi-bin/methods2010/material/stataCamp_3_ho.pdf`
`http://www.stanford.edu/~roymill/cgi-bin/methods2010/material/stataCamp_4_ho.pdf`
`http://www.stanford.edu/~roymill/cgi-bin/methods2010/material/stataCamp_5_ho.pdf`
`http://www.stanford.edu/~roymill/cgi-bin/methods2010/material/stataCamp_6_ho.pdf`
(but especially Roy Mill numbers 4 and 5).
2. Write a Stata program that computes pair matching.
3. Write a Stata/MATA program that computes pair matching. Compare the speed of the Stata program and the Stata/MATA program.

4. Write a Stata program that computes reweighting.
5. Write a Stata program that computes Blinder-Oaxaca.

Question C

1. Download the National Supported Work (NSW) demonstration data from Rajeev Dehejia's website: `http://www.nber.org/%7Erdehejia/nswdata.html`
2. Estimate the effect of the NSW program on annual earnings using pair matching, Blinder-Oaxaca, and reweighting.
3. Download `psmatch2.ado` using the command `ssc install psmatch2` and tinker with the matching estimators supported. Report the estimated treatment effect associated with matching on 5 neighbors and any other approach you find interesting.
4. Verify that your pair matching program agrees with `psmatch2`.
5. What do you think is the best *observational* estimate of the impact of the program on annual earnings? Explain your choice.
6. Do you believe that the overlap assumption is met in the NSW data? Why or why not and in what sense? (Note: This is a subtle question that deserves some thinking.)

References

Blinder, Alan, "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources*, 1973, *8* (4), 436–455.

Busso, Matias, John E. DiNardo, and Justin McCrary, "New Evidence on the Finite Sample Properties of Propensity Score Matching and Reweighting Estimators," 2009. Unpublished manuscript, UC Berkeley.

Dehejia, Rajeev H. and Sadek Wahba, "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs," *Journal of the American Statistical Association*, December 1999, *94* (448), 1053–1062.

DiNardo, John E., Nicole M. Fortin, and Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, September 1996, *64* (5), 1001–1044.

Imbens, Guido W., "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review," *Review of Economics and Statistics*, February 2004, *86* (1), 4–29.

Oaxaca, R., "Male-female wage differentials in urban labor markets," *International Economic Review*, 1973, *14*, 693–709.

Rosenbaum, Paul R. and Donald B. Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, April 1983, *70* (1), 41–55.

Smith, Jeff and Petra Todd, "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?," *Journal of Econometrics*, September 2005, *125* (1–2), 305–353.