

# Political Science 239

## Problem Set 6

Due date: Wednesday, October 18th, 2006

In this problem set, we will use the *non*-experimental data used by Lalonde to estimate the impact of the NSW employment program on earnings. This is, the treatment observations are the same that you used in the last problem set, but the control observations are not the randomized controls but rather observations from the CPS. You will be working with a random sample of size 4,000 from the original 16,000 non-experimental controls used by Lalonde. In order to use the dataset, once you've downloaded the file "lalonde2.RData", use the command `load(file=$path/lalonde2.RData)`. This will load an R workspace that has in it a dataset called `lalonde2`. This is the dataset that you should use. As always, you can use `names(lalonde2)`, `dim(lalonde2)`, etc., to check what's in this dataset. Just as in last problem set, the dependent variable is earnings in 1978 (`re78`), the treatment is whether the person was assigned to the training program (`treat`), and the rest are covariates. You can read the details of the experiment in Lalonde (1986) and Dehejia and Wahba (1999). This is a single-question problem set.

**Exercise 1** *The goal of this problem set is to find the best balance in this dataset. You should test the balance before matching, evaluate and explain whether you need to use matching (and why), and then try different matching procedures such as pscore matching, Mahalanobis matching, pscore plus Mahalanobis distance matching, genetic matching, regression adjustment, etc. You are free to try as many matching procedures as you like. After running your analysis, decide what method achieves the best balance and explain why this is the best balance that you can achieve. Use this*

*method to estimate the ATT and compare your answer with the answer you obtained in Problem Set 5. Remember, while you try different matching procedures, you should not be looking at the estimated ATT (i.e., do not define  $y$  in `Match()` until the end). Discuss your analysis and your results.*