PublPol 303D **Winter 2022**

Problem Set III

Due: 9am, Monday January 31st, 2022

Submit the answers in a document. Be concise but also clear what numbers you are reporting, and answer in full sentences. You should also hand in supporting code, but all the answers should be in a pdf or word file. For this assignment you should work in groups of 2 students.

Use the lottery.txt data. These are data from the 2001 paper with Rubin and Sacerdote. There are 496 observations on eight variables in this data set, in order, **earn** (annual earnings, averaged over six years after winning the lottery), **prize** (the annual prize won, calculated as the total prize divided by twenty), **lagged_earn** (average earnings over six years prior to winning the lottery), **age** (age in years), **male** (an indicator for being male), **educ** (years of education in years), **tixbot** (number of tickets bought per week), and **workthen** (an indicator for working at the time the lottery tickets were bought). All dollar amounts are measured in thousands of dollars.

We are interested in the effect of an additional dollar in yearly prize money on total labor earnings, known as the "marginal propensity to earn out of unearned income." This parameter is an important component in the evaluation of the disincentive effects of social welfare programs.

An experiment that would allow us to estimate that effect would be to take a set of individuals. Then select half of these individuals at random to receive $10,000 every year for the next ten years. Record their labor earnings during that period, and compare their average labor earnings to those of the control group that did not get the $10,000 every year. Divide the average difference by 10,000 to get the mpe.

Next we study the lottery data.

1. Read in the data and give the variables names and descriptive labels.

2. Create new variable "winner" that is equal to 1 for people with a positive prize and equal to zero for someone with a zero annual prize.

3. Regress average post-lottery earnings on the yearly prize and report the estimated mpe and the associated 95% confidence interval.

4. What are the reasons why you might not find this estimate credible?

5. Calculate the regression of each of the six pre-lottery variables on the annual prize. How do you interpret these correlations? Why are they not zero given that the lottery is (presumably) random?

6. To make the estimates of the mpe more credible, we consider two strategies. First, you may wish to limit your sample. Second, you may wish to include additional variables in the specification of your regression model. Choose your sample and specification and report the estimates of the mpe and the 95% confidence interval. Write one page defending your choices.

7. A policy maker is considering introducing a universal basic income guarantee for everybody of 20K per year, irrespective of how much they earn. The policy maker is concerned that this may reduce the incentives to work. Estimate the average effect on labor earnings of such a policy.

8. Suppose the universal income would only apply to people who currently make less than 40K per year. How would that change your answer? Incorporate use of interaction effects in your regression to address this question.