**Independent Practice**

Instructions: Submit this lab report to BlackBoard by 3PM on September 17. You should upload two files: one MS Word file with answers, and a Stata .do with replication code. In particular, questions must be properly numbered, and include graphs or tables as appropriate. Don't forget to include identifying information (Name, PUID, Section). The total possible point for this report is 10.

**After reading Gerber, Green and Larimer (GGL), we might expect that the group of people who received a treatment will have higher turnout than the group which received no treatment. The dataset contains variables called *voted* and *treatment*. (We use the word “treatment” because receiving a mailing is equivalent to a medical environment where a patient receives a pill or a procedure).**

1. What is dependent variable and the independent variable in the expectation above? Draw the causal diagram to represent the causal argument. (2pt)
2. Use the **.codebook** or **.tabulate** command to report the level of measurement for *voted* and *treatment*?(1pt)
3. Can you use these variables to compare how the two groups voted? Specifically, can you compare the group that received a treatment to the group that did not receive a treatment? Why or why not? (2pt)
4. Use the **.recode, generate** command from Pollock chapter 3 to create a new dummy variable called *treatment\_dummy* which gives subjects that received any of the treatments a value of 1, and subjects that did not received a treatment the value of 0. Be sure to include labels that say “received treatment” for values of 1, and “control” for values of 0. Please provide a code. (1pt)
5. Use **.codebook**, **.tabulate**,and/or **.summarize** to figure out the level of measurement of “treatment\_dummy” and to report its mean, median, and mode. What is the preferred measure of central tendency and why? (1pt)
6. We’re not quite ready to formally test statements like the one we state above, but we do have natural intuitions about what the outcome will look like if it is true. Which group would have a higher percentage of voters if it was true, the group that “received treatment” or the “control” group? (1pt)
7. Use the command **.tabulate** with the **if** option, and the variables *voted* and *treatment\_dummy* to report the percentage of each group that voted. Report this same result using the command .**histogram** with the **if** option. Do the results confirm your expectation from #6 above? (2pt)