GV300: Quantitative Political Analysis

Problem Set 10

Due Thursday, February 14, 9.45am on Faser

1. (25 marks) **Get out the vote - field experiment:**
   Download the paper "Partisan and Nonpartisan Message Content and Voter Mobilization: Field Experimental Evidence" by Costas Panagopoulos, it is available on Moodle.

   (a) (5 marks) What is the research question, the dependent variable (Y), the experimental manipulation (M), and the intended treatment (T)?

   (b) (10 marks) In an ideal experiment, the experimenter would manipulate the partisan identity of subjects directly (i.e., randomly assign subject i to be either Republican or Democrat). Why is this not possible? How is the author dealing with this issue in implementing the presented experiment, i.e. what is his manipulation?

   (c) (10 marks) Which observables is the author controlling for in this study and how is he implementing these controls in experimental design and statistical analysis?

2. (25 marks) **Instrumental variable I: Basics**

   (a) (5 marks) What does *exclusion restriction* mean in the context of instrumental variable regression?

   (b) (20 marks) Open your favorite statistical software and generate 500 observations. Generate variables `T` and `X` as a random draw from some distribution. Generate an outcome variable `Y` that is some function of `T` and `X`. Create an instrumental variable `Z`. Explain and demonstrate graphically why your variable `Z` is a perfect instrumental variable. Create a variable `zNot`, which is not a perfect instrument. How would such a variable need to look like? Explain and demonstrate graphically how `Z` and `zNot` differ making `zNot` an imperfect instrument.

3. (50 marks) **Instrumental variables II: Playing with Stata or R**

   (a) (5 Marks) Create a dataset with 50,000 observations. Create a variable called `Instrument`, which takes a value of 0 60% of the time and a value of 1 otherwise. Generate a variable `ObservableThing` $\sim N(0,1)$ and a variable `UnobservableThing` which is also $\sim N(0,1)$. Create the variable `VariableOfInterest`, which equals 1 if `ObservableThing + UnobservableThing + Instrument` ≥ 2.5 and which equals 0 otherwise. Finally, generate `DependentVariable = UnobservableThing + VariableOfInterest + e`, where `e` $\sim N(0,1)$. What is the causal effect of `VariableOfInterest` on `DependentVariable`?

   (b) (5 Marks) Regress `DependentVariable` on `VariableOfInterest` and interpret the coefficient. Why is this not close to the causal effect from part (a)?

   (c) (10 Marks) Use `Instrument` as the instrumental variable for `VariableOfInterest` in the regression of `DependentVariable` on `VariableOfInterest`. Calculate it both by hand and using the proper commands in Stata or R. Comment on what you observe.

```
# Here is what you need in R:
install.packages('AER')
```

```
# Define Y = outcome variable, T = treatment, Z = instrument, then:
ivreg(Y~T|Z)

// Here is the Stata command
// Define Y = outcome variable, T = treatment, Z = instrument, then
ivregress 2sls Y (T=Z)
```

(d) (10 Marks) Change the setup from part (a) such that `Instrument` is only equal to 1 2% of the time. Then rerun the regression in part (c) (using Stata or R) and comment on what happens. What's wrong with this instrument?

(e) (10 Marks) Change the setup from part (a) such that `VariableOfInterest = ObservableThing + UnobservableThing + 0.05*Instrument`. Then rerun the regression in part (c) (using Stata or R) and comment on what happens.

(f) (10 Marks) Change the setup from part (a) such that `DependentVariable = UnobservableThing + VariableOfInterest + Instrument + e`. Then rerun the regression in part (c) (using Stata or R) and comment on what happens.