Intro Econometrics[1]

Lab 10: Dummy Variables

Note: This lab is an important part of your homework! Whatever part you don't finish with your TAs, it is homework. We will post a stata file that runs everything and provides the answers later.

Helpful formula: If r, s, are constants and X and Y RVs, then

$$VAR(rX + sY) = r^2 VAR(X) + s^2 VAR(Y) + 2rs COV(X, Y).$$

Models with single dummies

One of the most "famous" economics studies was performed by David Card and Alan Krueger (CK), published in the American Economic Review in 1994. They studied how the fast food labor market in New Jersey and Pennsylvania was affected by an increase in the minimum wage that occurred in New Jersey in 1992. Using data from 1991 and 1992, they found the increase in minimum wage did not reduce employment – it *increased* employment.

Using our data on minimum wages again from lab 2, we will see what we can find regarding this NJ minimum wage increase.

1. Load the lab2minwagedata.dta dataset into stata. You can find it in the Lab 3 folder on Canvas under Lab Materials.

2. Limit the data to only years 1991 and 1992. Also create variables containing the natural log of epteen and minwage.

3. Create a dummy variable to identify new jersey called "NJ" where

$$NJ = \begin{cases} 1 & \text{if state} = 34 \\ 0 & \text{otherwise.} \end{cases}$$

   The variable state contains numeric codes to identify states, and 34 identifies New Jersey. Pennsylvania is identified by code 42.

4. Next, estimate the following model: $\ln(epteen) = \beta_0 + \beta_1 \ln(minwage) + u$ on only observations from New Jersey and Pennsylvania for the years 1991 and 1992. Since your

---

[1]Clemson University, Scott Barkowski, Instructor.

data is already limited to the relevant years, all you need to do is add an "if" statement to your regression command which will limit it to only the observations from the correct states. One way to do this would be to use the "inlist" function. What sort of estimate do you get for the effect of minimum wage on employment? Is it consistent with what CK found in their study?

5. Next, allow the model to have two intercepts by adding your NJ dummy to the model (continue limiting to only NJ and PA). What happens to the result? How does it compare to CK's result? Use paper and pencil to plot the estimated conditional expectation functions for New Jersey and Pennsylvania. Be sure to note the important features (intercepts, slope).

6. How would you explain what the estimate on the NJ dummy means to someone who doesn't know what intercepts are?

7. Next create a variable representing the natural log of minimum wage interacted with the NJ dummy. That is, $lmwNJ = \ln(minwage) \times NJ$.

8. Now use OLS to estimate the model

$$\ln(epteen) = \beta_0 + \beta_1 \ln(minwage) + \beta_2 NJ + \beta_3 \ln(minwage) \times NJ + u,$$

but, of course, use the variable $lmwNJ$ as the interaction term in your regression. Also, continue limiting to only NJ and PA.

9. Use the results from above to plot (on paper) the estimated conditional expectation functions, noting again the key features. Are their important differences from the previous plot you made when there wasn't a interaction term?

10. Test the hypothesis that the slopes of the two CEFs are different. Choose your significance level, but be sure to choose your critical value using the correct number of degrees of freedom.

11. Next test the hypothesis that the minimum wage elasticity of employment for teenagers in New Jersey is equal to zero. That is, that the slope of the New Jersey function is equal to zero. To do this, you will need to find a covariance term between two estimators. See question 1 from Lab 9 for how to get stata to give you that. You can choose your own significance level, but be sure to use the correct number of degrees of freedom to find your critical value.

12. EXTRA CHALLENGE: You can also test the hypothesis above by rearranging the variables in the regression in the manner we used last week. Can you figure out how to do it? Hint: add and subtract $\beta_3 ln(minwage)$ from your regression.

Models with multiple dummies

We can also use dummies when we have more than two groups. So far we limited ourselves to only New Jersey and Pennsylvania – two groups. But let's open it to all states.

13. Estimate the model from question 4 without restricting to NJ and PA. What does this say about the effect of minimum wages?

14. Now add a dummy for every state and DC (except one). This would require you to create a lot of dummy variables, but fortunately Stata makes it easier by having a way you can tell it to automatically create those dummies for you. You can tell stata to add a dummy for every state/dc in the data (except one) by adding the following to your regression command, "i.state", as if it were a control variable. Does doing this change your result from the previous regression?

15. Test the hypothesis that state is statistically significant. That is, that state dummies jointly have no effect on employment. Choose your own significance level.