ECON 6511: Advanced Applied Econometrics
Homework 3
Due in class January 31, 2018

1. (Wooldridge, Chapter 13, Problem 3) Use the data in KIELMC.dta for this exercise.

   (a) The variable *dist* is the distance from each home to the incinerator site, in feet. Consider the model:

   $$\log(price) = \beta_0 + \delta_0 y81 + \beta_1 \log(dist) + \delta_1 y81 \cdot \log(dist) + u$$

   If building the incinerator reduces the value of homes closer to the site, what is the sign of δ_1? What does it mean if $\beta_1 > 0$?

   (b) Estimate the model from part (a) and report the results in equation form. Interpret the coefficient on $y81 \cdot \log(dist)$. What do you conclude?

   (c) Add age, age^2, $rooms$, $baths$, $\log(instst)$, $\log(land)$, and $\log(area)$ to the equation. Now, what do you conclude about the effect of the incinerator on housing values?

   (d) Why is the coefficient on $\log(dist)$ positive and statistically significant in part (b) but not in part (c)? What does this say about the controls used in part (c)?

2. (Wooldridge, Chapter 13, Problem 5) Use the data in RENTAL.dta for this exercise. The data for the years 1980 and 1990 include rental prices and other variables for college towns. The idea is to see whether a stronger presence of students affects rental rates. The unobserved effects model is:

   $$\log(rent_{it}) = \beta_0 + \delta_0 y90_t + \beta_1 \log(pop_{it}) + \beta_2 \log(avginc_{it}) + \beta_3 pctstu_{it} + a_i + u_{it}$$

   where *pop* is city population, *avginc* is average income, and *pctstu* is student population as a percentage of city population (during the school year).

   (a) Estimate the equation by pooled OLS and report the results in equation form. What do you make of the estimate on the 1990 dummy variable? What do you get for $\hat{\beta}_{pctstu}$?

   (b) Are the standard errors you report in part (a) valid? Explain.

   (c) Now, difference the equation and estimate by OLS. Compare your estimate of β_{pctstu} with that from part (b). Does the relative size of the student population appear to affect rental prices?

   (d) Obtain the heteroskedasticity-robust standard errors for the first-differenced equation in part (c). Does this change your conclusions?

3. (Wooldridge, Chapter 13, Problem 14) Use the data in JTRAIN3.dta for this question.

(a) Estimate the simple regression model $re78 = \beta_0 + \beta_1 train + u$, and report the results in equation form. Based on this regression, does it appear that job training, which took place in 1976 and 1977, had a positive effect on real labor earnings in 1978?

(b) Now use the change in real labor earnings, $cre = re78 - re75$, as the dependent variable. (We need not difference *train* because we assume there was no job training prior to 1975. That is, if we define $ctrain = train78 - train75$ then $ctrain = train78$ because $train75 = 0$.) Now what is the estimated effect of training? Discuss how it compares with the estimate in part (a).

(c) Find the 95% confidence interval for the training effect using the usual OLS standard error and the heteroskedasticity-robust standard error, and describe your findings.

4. (Wooldridge, Chapter 13, Problem 8) VOTE2.data includes panel data on House of Representatives elections in 1988 an 1990. Only winners from 1988 who are also running in 1990 appear in the sample; these are the incumbents. An unobserved effects model explaining the share of the incumbent's vote in terms of expenditures by both candidates is:

$$vote_{it} = \beta_0 + \delta_0 d90_t + \beta_1 \log(inexp_{it}) + \beta_2 \log(chexp_{it}) + \beta_3 incshr_{it} + a_i + u_{it},$$

where $incshr_{it}$ is the incumbent's share of total campaign spending (in percentage form). The unobserved effect a_i contains characteristics of the incumbent–such as "quality"–as well as things about the district that are constant. The incumbent's gender and party are constant over time, so these are subsumed in a_i. We are interested in the effect of campaign expenditures on election outcomes.

(a) Difference the given equation across the two years and estimate the differenced equation by OLS. Which variables are individually significant at the 5% level against a two-sided alternative?

(b) In the equation from part (a), test for joint significance of $\Delta\log(inexp)$ and $\Delta\log(chexp)$. Compute by hand and report the F-statistic.

(c) Reestimate the equation from part (a) using $\Delta incshr$ as the only independent variable. Interpret the coefficient on $\Delta incshr$. For example, if the incumbent's share of spending increases by 10 percentage points, how is this predicted to affect the incumbent's share of the vote?

(d) Redo part (c), but now use only pairs that repeat challengers. [This allows us to control for characteristics of challengers as well, which would be in a_i.]