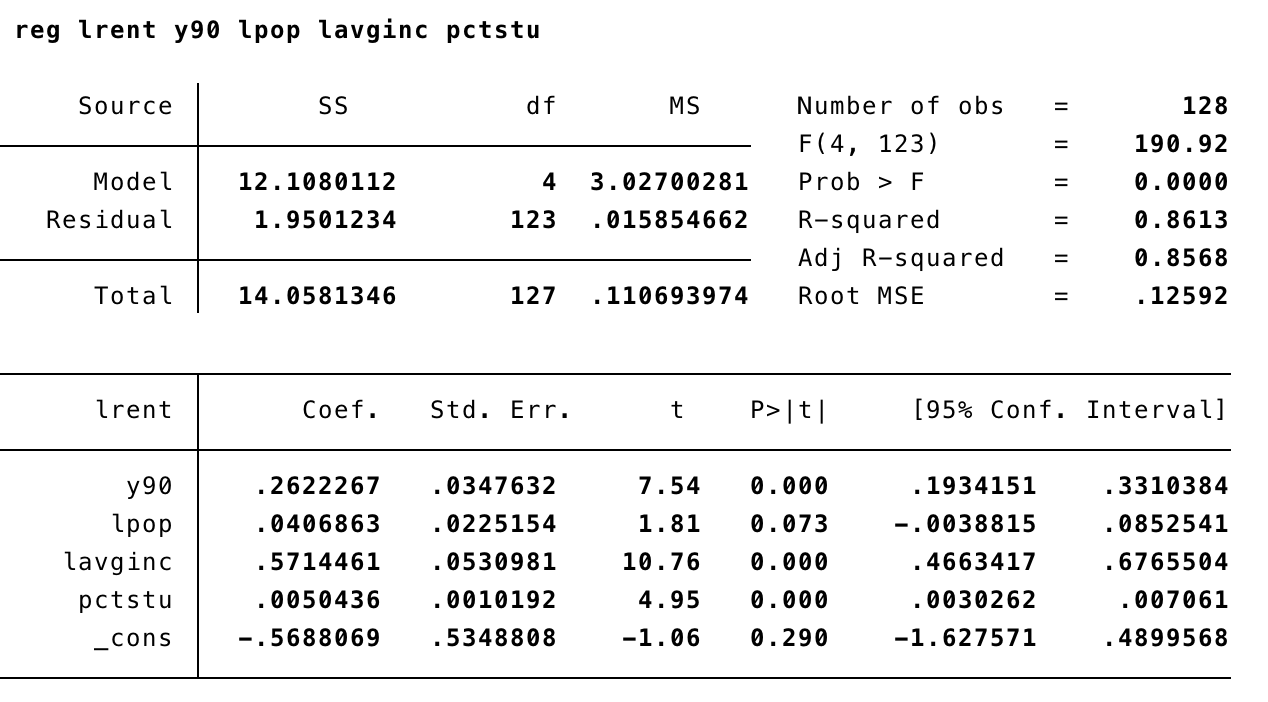
ECON 6511: Advanced Applied Econometrics

Homework 4

February 7, 2017

Surabhi Asati

1. (Wooldridge, Chapter 14, 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(rentit) = β0 + δ0y90t + β1log(popit) + β2log(avgincit) + β3pctstuit + ai + uit where pop is city population, avginc is average income, and pctstu is student population as  a percentage of city population (during the school year).
2. 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 βˆpctstu?



log(rentit) = -0.569 + 0.262 y90t + 0.041 log(popit) + 0.571 log(avgincit) + 0.005 pctstuit + ai + uit

The 1990 dummy variable has a coefficient = 0.262 which means the rental rate has grown by 26.2 % from 1980 to 1990 keeping other variables unchanged.

The student population as a percentage of city population has coefficient = 0.005 which means 1% increase in student population increases rent by 0.5%

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

The standard errors in part (a) is not valid because ai is not present

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

Estimated equation:

∆log(rentit) = 0.386 + 0.072 y90t + 0.041 ∆log(popit) + 0.310 log(avgincit) + 0.112 ∆pctstuit + uit

the coefficient of pctstu is now 0.112 which is way too high than it was in part (b). So, with 1% increase in pctstu, the rent will increase by 1.1%. It appears that the relative size of the student population affects the rental prices.

1. Estimate the model by fixed effects to verify that you get identical estimates and stan- dard errors to those in part (c).

Estimated equation:

log(rentit) = 0.386 y90t + 0.072 log(popit) + 0.310 log(avgincit) + 0.112 cpctstuit

coefficient for pctstu for both equations is 0.011 with p-value = 0.009and se =0.0041

1. (Wooldridge, Chapter 14, Problem 4) Papke (1994) studied the effect of the Indiana enterprise zone (EZ) program on unemployment claims. The author uses a model that allows each city to have its own time trend:  log(uclmsit) = ai + cit + β1ezit + uit, where ai and ci are both unobserved effects. This allows for more heterogeneity across cities.
2. Show that, when the previous equation is first differenced, we obtain ∆log(uclmsit)=ci +β1∆ezit +∆uit, t=2,...,T.  Notice that the differenced equation contains a fixed effect, ci.

Given:

log(uclmsit) = ai + cit + β1ezit + uit ----1st equation

For each i, average value over time (^):

log(uclmsi^)=β1 ezi^ + ai + ci + ui^, t=2,...,T ……2nd equation

subtract equation 2nd from 1st , we get:

∆log(uclmsit) = β1∆ezit + ci +∆uit, t=2,...,T

1. Estimate the differenced equation by fixed effects using the data in EZUNEM.DTA. What is the estimate of β1? Is it very different from the estimate in textbook Example 13.8 of −0.182? Is the effect of enterprise zones still statistically significant?

Estimated equation:

∆log(uclmsit) = -0.252∆ezit + ci +∆uit,

Estimate of β1 = -0.211664 which is larger than – 0.182 by 0.70. Unemployment falls by 22.2%. β1 is statistically significant with p-value = 0.028

1. Add a full set of year dummies to the estimation in part (b). What happens to the estimate of β1?

Estimate of β1 = -0.2511664. P-value = 0.017 at 5% significance. So, employement rate decreases to 17.5% [exp(-0.01919402)-1]

This is smaller than estimates in part (b)

3. (Wooldridge, Chapter 14, Problem 7) Use the state-level data on murder rates and executions in MURDER.dta for the following exercise

1. Consider the unobserved model mrdrteit = ηt + β1execit + β2unemit + ai + uit,  where ηt simply denotes different year intercepts and ai is the unobserved state effects. If past executions of convicted murderers have a deterrent effect, what should be the sign of β1? What sign do you think β2 should have? Explain.

If past executions of convicted murderers have a deterrent effect, β1 has negative sign as β1 < 0.

Β2 will be positive , β2 > 0 Murder rate will increase due to unemployment

1. Using just the years 1990 and 1993, estimate the equation from part (a) by pooled OLS. Ignore the serial correlation problem in the composite errors. Do you find any evidence for a deterrent effect?

β1 = 0.1277. No evidence of deterrent effect. This shows that increase in execution by 1 will increase murder by 0.1277. with p-value = 0.63, it is not statistically significant.

1. Now, using 1990 and 1993, estimate the equation by fixed effects. You may use first differencing since you are only using two years of data. Is there evidence of a deterrent effect? How strong?

β1 = -0.131 with p-value = 0.375. So it not significant. There should be 10 executions for a deterrent effect. 0.1 \* 10 = 1 - execution is low. So this is moderately strong.

1. Compute the heteroskedasticity-robust standard error for the estimation in part (c).

heteroskedasticity-robust standard error has come down to 0.1331 from 0.1473

1. Find the state that has the largest number for the execution variable in 1993. (The variable exec is total executions in 1991, 1992, and 1993.) How much bigger is this value than the next highest value?

state that has the largest number for the execution variable in 1993 = TX

Next highest = VA

1. Estimate the equation using first differencing, dropping Texas from the analysis. Com- pute the usual and heteroskedasticity-robust standard errors. Now, what do you find? What is going on?

1. Use all three years of data and estimate the model by fixed effects. Include Texas in the analysis. Discuss the size and statistical significance of the deterrent effect compared with only using 1990 and 1993.

4. (Wooldridge, Chapter 14, Problem 14) Use the data set in AIRFARE.dta to answer this question. The estimates can be compared with those at the end of the last lecture.

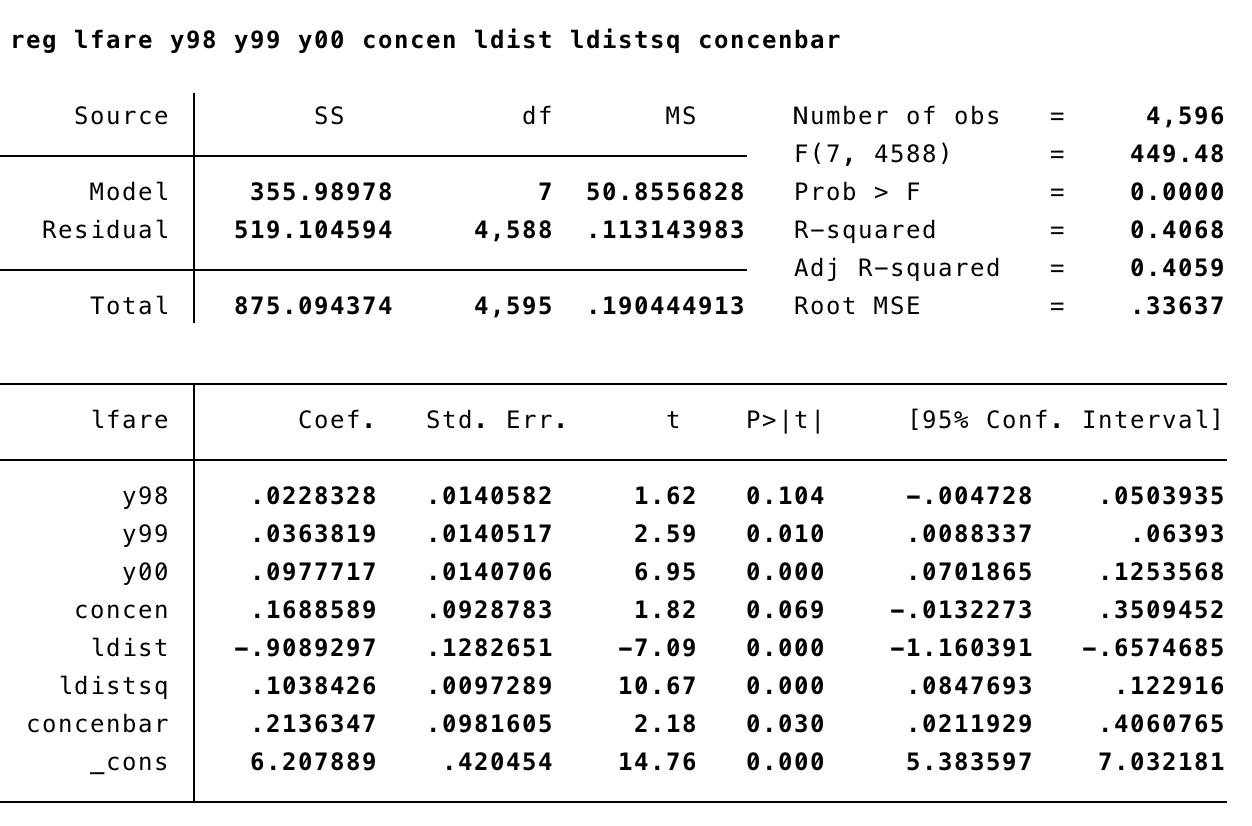
1. Compute the average of the variable concen for each route; call these concenbar. How many different time averages can there be? Report the smallest and the largest. Note: You can generate a variable containing the average value of concen for each route using the Stata command: “egen concenbar = mean(concen), by(id)”

Concenbar = 1149

smallest = 0.1862

largest = 0.9997

1. Estimate the equation: lfareit =β0 + δ1y98t + δ2y99t + δ3y00t + β1concenit + β2ldistt + β3ldistsqi  + γ1concenbari + ai + uit by random effects. Verify that βˆ1 is identical to the FE estimate computed in class.



lfareit = 6.207 + 0.022 y98t + 0.036 y99t + 0.097 y00t + 0.168 concenit – 0.908 ldistt + 0.103 ldistsqi  + 0.213 concenbari + uit

β1 is 0.1688589. This is identical to the FE estimate computed in class.

1. Using the equation from part (b) and the usual RE standard error, test H0 : γ1 = 0 against the two-sided alternative. Report the p-value. What do you conclude about RE versus FE for estimating β1 in this application?

H0: Y1 = 0

H1: y1 != 0

Two-tailed p-value for Y1 = 0.002

Se = 0.0678

for estimating β1 in this application, RE and FE give same values. This indicates lots of variation in ai

1. Obtain a t-statistic (and, therefore, p-value) that is robust to arbitrary serial correlation and heteroskedasticity using “, cluster(id)”. Does this change the conclusion reached in part (c)?

Concen = 0.169. This does not change the conclusion reached in part (c)