Homework 4 – due 11:59 pm, April 14th

(Please submit a single pdf file)

1. Redo the example for job training grants and firm scrap rates (on page 31 of the lecture notes from week 8) by taking the firm level heterogeneity into consideration. Remember, in the class we did the fixed effect model and random effect model to incoorporate the firm level unobserved effect.

a. Include the lag of log($scrap$) ($lscrap\_1$) in your model and calculate the OLS estimator. Your model is

$$lscrap = \beta_1 + \beta_2 I(year = 1988) + \beta_3 I(year = 1989) + \beta_4 grant + \beta_5 grant\_1 + \beta_6 lscrap\_1 + u.$$

b. Compare the results obtained from part a) with what we get in class from a fixed efffect model. What is the difference between the two versions of $\hat{\beta}_4$ and $\hat{\beta}_5$? Why is there such a difference?

c. Include the lag of log($scrap$) ($lscrap\_1$) in your model and calculate $\underline{\hat{\beta}}$ using the random effect model:

$$lscrap = \beta_1 + \beta_2 I(year = 1988) + \beta_3 I(year = 1989) + \beta_4 grant + \beta_5 grant\_1 + \beta_6 lscrap\_1 + c + u.$$

d. After running a test for $H_0 : \sigma_c^2 = 0$ vs. $H_0 : \sigma_c^2 \neq 0$. We get a test statistic $T = 1.87$. We know that for a random variable Z following a standard normal distribution, $P(Z < 1.645) = 0.95$, $P(Z < 1.95) = 0.975$ and $P(Z < 2.33) = 0.99$. Will you reject the null hypothese at $\alpha = 0.05$? Based on your hypothesis testing results, which model should we use – model specified in part a or part c?

2. Analyze the DoctorVisits data (available in "AER" package), taken from Cameron and Trivedi (1998), using a Poisson regression for the number of visits. Is the Possion model satisfactory? If not, where are the problems and what could be done about them?

3. Analyze the Affairs data (available in "AER" package) utilizing models for count data. Fit the following 4 models: a) Poisson regression b) Negative Binomial c) Hurdle Poisson d) Hurdle Negative binomial discuss their performance by comparing: 1) Log likelihood; 2) AIC; 3) Predicted vs. actual (round(E(y)) vs. y); 4) rootgram.