ECON 510

Empirical Homework

1) Consider following linear regression model:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i \ \ \forall i = 1, 2, ..., 253$$

a) Report the OLS estimation results
b) Comment on individual significance of each parameter.
c) Test overall significance of the model. Carefully state each step of the hypothesis test.
d) (Bonus)Test the hypothesis that at least one of the parameters is 0. Write down the null in statistical form and perform the test without using any other external knowledge or guess.
e) Find the redundant regressors by using goodness of fit measures if you believe there exist one.
f) Repeat the exercise in f with a F test. Clearly, demonstrate the test procedure. Propose the best model you find.
g) Test the normality of the residuals in final model you propose. Check if residuals contain any problems that violate classical regression assumptions. Properly test the problem you suspect.

2) We have a panel data described in excel file.
   a) Construct a econometric model with these variables.
   b) Estimated the model as if there is no unobserved heterogeneity and report the results, comment on significance of the model and individual significance.
   c) Graphically check if residuals have same mean across airlines. Comment on results. Do the same analysis for the time.
   d) Run a Fixed effect with period, cross-section and both period and cross-section. Informally, comment which model can be used.
   e) Test the fixed effect with your favorite method.
   f) Run a random cross-section effect (without any fixed effect on period or cross-section.
   g) Test random effect versus fixed effect.

3) Consider following simple regression model:

$$y_i = 1.5 + 2.5 x_i + \varepsilon_i$$

where x_i's are random iid draws from Uniform distribution [10,12] and ε_i are iid standard normally distributed.

Simulate the above data 1000 times (with different x_i and ε_i sequences in each simulation) with sample size 10, 100 and 1000 (with your favorite statistical computer program). For each simulation compute OLS coefficients and s^2 .

Comment on the results. What can you say about consistency and asymptotic normality?