JS ECONOMETRICS MODULE I – C. NEWMAN
TOPIC 3: MULTIPLE REGRESSION ANALYSIS

HOMEWORK 3

PLEASE ANSWER ALL QUESTIONS.
PLEASE SUBMIT SOLUTIONS AT THE TUTORIAL ON WEDNESDAY NOVEMBER 18TH,
ROOM 2041A, 6PM

1. The median starting salary for new law graduates is determined by:

   $$\log(salary) = \beta_0 + \beta_1 SAT + \beta_2 GPA + \beta_3 \log(libvol) + \beta_4 \log(cst) + \beta_5 rank + u$$

   where
   SAT is the median SAT score for the graduating class (higher the better)
   GPA is the median college Grade Point Average for the class (higher the better)
   libvol is the number of volumes in the law school library
   cst is the annual cost of attending law school
   rank is the law school ranking with rank=1 being the best

   (i) Explain why we expect $\beta_5 \leq 0$
   (ii) What signs do you expect for the other slope parameters? Explain.
   (iii) The estimated equation is:

   $$\log(\hat{salary}) = 8.34 + 0.005SAT + 0.25GPA + 0.09\log(libvol) + 0.04\log(cst) - 0.003rank$$

   $$n = 136 \qquad R^2 = 0.842$$

   Interpret the estimated coefficients on the variables in this model and comment on the reported R^2

   (20 marks)

2. Suppose you estimate the following equation using data on working men:

   $$\hat{educ} = 10.36 - 0.094sibs + 0.131meduc + 0.210feduc$$

   $$n = 722 \qquad R^2 = 0.214$$

   where
   educ is years of schooling
   sibs is the number of siblings
   meduc is mother's years of education
   feduc is father's years of education

   (i) Does *sibs* have the expected effect? Explain. Holding *meduc* and *feduc* constant, by how much does *sibs* have to increase to reduce predicted years of education by one year?
   (ii) Discuss the interpretation of the coefficient on *meduc*.
   (iii) Suppose Man A has no siblings and his mother and father each have 12 years of education. Suppose Man B has no siblings and his mother and father have 16 years of education. What is the predicted difference in years of education between B and A?

   (15 marks)

3. Consider the following model:

$Y_i = \beta_0 + \beta_1 Educ_i + \beta_2 Exper_i + u_i$

(i) Explain how you would estimate this model using Ordinary Least Squares

(ii) What assumptions are required to show that the OLS estimators are unbiased and efficient? Say that all individuals surveyed are 50 years old and you construct the variable *Exper* as $Exper_i = 50 - Educ_i - 4$. Can this model be estimated using OLS?

(iii) Suppose instead of estimating this model you estimate

$Y_i = \alpha_0 + \alpha_1 Educ_i + u_i$ using OLS.

What is the relationship between $\hat{\alpha}_1$ and $\hat{\beta}_1$?

Comment on the properties of $\hat{\alpha}_1$.

(iv) Under what circumstances would $\hat{\alpha}_1 = \hat{\beta}_1$?

(40 marks)

4. Suppose that you are interested in estimating the ceteris paribus relationship between *Y* and X_1. You collect data on two control variables, X_2 and X_3. Let $\tilde{\beta}_1$ be the simple regression estimate from *Y* on X_1 and $\hat{\beta}_1$ be the multiple regression estimate from *Y* on X_1, X_2 and X_3.

(i) If X_1 is highly correlated with X_2 and X_3 in the sample, and X_2 and X_3 have large partial effects on *Y*, would you expect $\tilde{\beta}_1$ and $\hat{\beta}_1$ to be similar or different? Explain

(ii) If X_1 is almost uncorrelated with X_2 and X_3, but X_2 and X_3 are highly correlated, will $\tilde{\beta}_1$ and $\hat{\beta}_1$ to be similar or different? Explain

(iii) If X_1 is highly correlated with X_2 and X_3, and X_2 and X_3 have small partial effects on *Y*, would you expect $Var(\tilde{\beta}_1)$ or $Var(\hat{\beta}_1)$ to be smaller? Explain.

(iv) If X_1 is almost uncorrelated with X_2 and X_3, X_2 and X_3 have large partial effects on *Y*, and X_2 and X_3 are highly correlated, would you expect $Var(\tilde{\beta}_1)$ or $Var(\hat{\beta}_1)$ to be smaller? Explain.

(25 marks)