ECON 6306 Fall 2016
Applied Econometrics Problem Set 2

1. Does the return to education differ by race and gender? For this exercise, use the file *cps4.csv*. In this exercise, you will extract subsamples of observations consisting of (i) all males, (ii) all females, (iii) all whites, (iv) all blacks, (v) white males, (vi) white females, (vii) black males, and (viii) black females.

   (a) For each sample partition, obtain the summary statistics of wage.

   (b) For each sample partition, estimate the log-linear model

   $$ln(WAGE) = \beta_1 + \beta_2 EDUC + e$$

   How do you interpret the coefficient on Education for each group?

   (c) Does the model fit the data equally well for each subsamples?

   (d) For each sample partition, test the null hypothesis that rate of return to education is 10% against the alternative that it is not, using a two-tail test at the 5% level of significance.

2. This question is concerned with the value of houses in towns surrounding Boston. It uses data of Harrison, D., and D.L. Rubindfeld (1978)," Hedonic Prices and the Clean Air," *Journal of Environment Economics and Management*, 5, 81-102. The output appears in Table below. The variables are defined as follows.
   VALUE = Median value of owner-occupied homes in $1000's
   CRIME = per capita crime rate by town
   NITOX = nitric oxides concentration (parts per 10 million)
   ROOMS = average number of rooms per dwelling
   AGE = proportion of owner-occupied units built prior to 1940
   DIST = weighted distances to five Boston employment centres
   ACCESS = index of accessibility to radial highways
   TAX = full-value property-tax rate per $10,000
   PTRATIO = pupil-teacher ratio by town

Dependent Variable: VALUE	Included Observations: 506			
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	28.4067	5.3659	5.29	0.0000
CRIME	-0.1834	0.0365	-5.03	0.0000
NITOX	-22.8109	4.1607	-5.48	0.0000
ROOMS	6.3715	0.3924	16.24	0.0000
AGE	-0.0477	0.0141	-3.39	0.0008
DIST	-1.3353	0.2001	-6.67	0.0000
ACCESS	0.2723	0.0723	3.77	0.0002
TAX	-0.0126	0.0038	-3.34	0.0009
PTRATIO	-1.1768	0.1394	-8.44	0.0000

   (a) Report briefly on how each of the variables influences the value of a home.

(b) Find 95% interval estimates for the coefficients of CRIME and ACCESS.

(c) Test the hypothesis that increasing the number of rooms by one increases the value of a house by $7,000.

(d) Test as an alternative hypothesis H_1 that reducing the pupil-teacher ratio by 10 will increase the value of a house by more than $10,000.

3. The file *br2.csv* contains data on 1080 houses sold in Baton Rouge, Louisiana, during mid-2005. We will be concerned with the selling price ($PRICE$), the size of the house in square feet ($SQFT$), and the age of the house in years (AGE).

(a) Use all observations to estimate the following regression model and report the results

$$PRICE = \beta_1 + \beta_2 SQFT + \beta_3 AGE + e$$

i. Interpret the coefficients

ii. Find a 95% interval estimate for the price increase for an extra square foot of living space- that is, $\delta PRICE/\delta SQFT$.

iii. Test the hypothesis that having a house a year older decreases price by 1000 or less ($H_0 : \beta_3 \geq -1000$) against the alternative that it decreases price by more than 1000($H_1 : \beta_3 < -1000$)

(b) Add the variables $SQFT^2$ and AGE^2 to the model in part(a) and re-estimate the equation. Report the results.

i. Find estimates of the marginal effect $\delta PRICE/\delta SQFT$ for the smallest house in the sample, the largest house in the sample, and a house with 2300 $SQFT$. Comment on these values. Are they realistic?

ii. Find estimates of the marginal effect $\delta PRICE/\delta AGE$ for the oldest house in the sample, the newest house in the sample, and a house that is 20 years old. Comment on these values. Are they realistic?

iii. Find a 95% interval estimate for the marginal effect $\delta PRICE/\delta SQFT$ for a house with 2300 square feet.

iv. For a house that is 20 years old, test the hypothesis

$$H_0 : \frac{\delta PRICE}{\delta AGE} \geq -1000$$

against

$$H_1 : \frac{\delta PRICE}{\delta AGE} < -1000$$

(c) Add the interaction variable $SQFT \times AGE$ to the model in part(b) and reestimate the equation. Report the results. Repeat parts(i),(ii),(iii), and (iv) from part (b) for this new model. Use $SQFT$=2300 and AGE=20.

(d) From your answers to parts (a),(b), (c), comment on the sensitivity of the results to the model specification.

4. In problem 2, an equation used for the valuation of homes in towns surrounding Boston was estimated. Data used for estimation is given by file *boston.csv*.

   (a) Reestimate that equation with White's standard errors and the report the results in a table.

   (b) For the coefficients of *CRIME,ROOMS,AGE,TAX*, compare 95% confidence intervals obtained using the standard errors from problem 3 with that of the current problem.

   (c) Do you think heteroskedasticity is likely to be a problem?

   (d) What misleading inferences are likely if the incorrect standard errors are used?