Econometrics Test
2015 - 06 - 03

Name: ______________________________

Matricola: ______________________ email: ______________________

1. Say if the following statements are unambiguously true (TRUE), unambiguously false (FALSE) or impossible to classify the way they are stated (CAN'T SAY). Write the motivations to your answers **only** in the space provided. A "CAN'T SAY" answer with no motivations will be considered wrong.

   (a) Covariance matrices are always symmetric.
   TRUE ◯ FALSE ◯ CAN'T SAY ◯

   (b) Let X be a continuous random variable with support $(-\infty, \infty)$. Then $E(X^2) > 0$.
   TRUE ◯ FALSE ◯ CAN'T SAY ◯

   (c) A consistent estimator may not have a limit in probability.
   TRUE ◯ FALSE ◯ CAN'T SAY ◯

   (d) In the dynamic model $y_t = 0.8y_{t-4} + 0.5x_t - 0.3x_{t-4}$ the dynamic multiplier of order 1, δ_1, equals 0.4
   TRUE ◯ FALSE ◯ CAN'T SAY ◯

   (e) In the dynamic model $y_t = 0.8y_{t-4} + 0.5x_t - 0.3x_{t-4}$ the long-run multiplier, c, equals 1.
   TRUE ◯ FALSE ◯ CAN'T SAY ◯

2. Suppose that you have two distinct datasets (call them A and B respectively), in which you observe the same variables. So you have a vector $\mathbf{y}_a$ with n_a elements and a corresponding $\mathbf{X}_a$ matrix with n_a rows and k columns, together with a vector $\mathbf{y}_b$ with n_b elements and a $\mathbf{X}_b$ matrix with n_b rows and k columns.

   You run OLS on the two datasets separately and obtain $\hat{\beta}_a$, $SSR_a = \mathbf{e}_a'\mathbf{e}_a$ for dataset A and $\hat{\beta}_b$, $SSR_b = \mathbf{e}_b'\mathbf{e}_b$ for dataset B.

   Then, you join the two subsamples and run OLS using the vector $\mathbf{y} = [\mathbf{y}_a' \quad \mathbf{y}_b']'$ and the matrix $\mathbf{X} = [\mathbf{X}_a' \quad \mathbf{X}_b']'$. Call the corresponding OLS statistic $\hat{\beta}$ and the sum of squared residuals SSR.

   Prove analytically that $SSR \geq SSR_a + SSR_b$. (*Hint: start by showing that* $SSR = SSR_a + SSR_b$ *if* $\hat{\beta}_a = \hat{\beta}_b$.)

3. Table 1 contains a *hedonic model*, that is a model in which we try to explain the price of a good on the basis of its characteristics. In this case, the dependent variable is the sale price of houses. We have a sample of 1080 houses in Baton Rouge, Louisiana, USA, for which the following variables are observed:

Variable name	Description
l_price	log of sale price, dollars
l_sqft	log of total square feet
age	age in years
bedrooms	number of bedrooms
baths	number of full baths
owner	1 if owner occupied at sale; 0 if vacant or tenant
pool	1 if pool present
waterfront	1 if on waterfront

   Answer the following questions on a separate sheet, making use of the numerical estimates that you find in table 1:

   (a) Consider a house 1600 square feet big, built 20 years ago with three bedrooms, two bathrooms, occupied at sale, with no pool and on the waterfront. Calculate the value of the house.[1]

   (b) Do we have a heteroskedasticity problem in this model? Motivate your answer.

   (c) Someone looked at these estimates and said "It looks as if smaller houses sell for a higher price per square foot". Is this statement correct?

   (d) Someone else looked at these estimates and said "Hmm. Apparently, the more bedrooms a house has, the lower its value is. This looks strange". Is this statement correct? Discuss.

   (e) A newspaper titled "In Baton Rouge, property depreciates by 0.5% a year". Is this claim supported by the data?

[1] Note: $\ln(1600) = 7.3778$.

```
Model 1: OLS, using observations 1-1080
Dependent variable: l_price
Heteroskedasticity-robust standard errors, variant HC1

               coefficient   std. error    t-ratio    p-value
  ------------------------------------------------------------
  const         5.94529      0.249874      23.79     8.18e-101 ***
  l_sqft        0.734324     0.0392789     18.70     1.00e-67  ***
  age          -0.00566465   0.000756504   -7.488    1.46e-13  ***
  bedrooms     -0.0542742    0.0187356     -2.897    0.0038    ***
  baths         0.228085     0.0262221      8.698    1.25e-17  ***
  owner         0.0687799    0.0171837      4.003    6.70e-05  ***
  pool          0.0634803    0.0337406      1.881    0.0602    *
  waterfront    0.186096     0.0363630      5.118    3.66e-07  ***

Mean dependent var   11.79518   S.D. dependent var   0.524535
Sum squared resid    84.83602   S.E. of regression   0.281315
R-squared            0.714234   Adjusted R-squared   0.712368
F(7, 1072)           205.5012   P-value(F)           4.1e-193

hite's test for heteroskedasticity:
Test statistic: TR^2 = 282.730411, with p-value = 0.000000

Restriction:
 b[l_sqft] = 1
Test statistic: 45.7496, with p-value = 2.20523e-11

Restriction:
 b[age] = -0.005
Test statistic: 0.771907, with p-value = 0.379825
```

Table 1: Hedonic model for house prices