B.E. International Program

Faculty of Economics, Thammasat University

Semester: 2/2014
EE325 Introductory Econometrics
Homework#3 (Due on April 23rd, 2015)

For Question 1-4, attempt only 2 questions

For Question 5-8, attempt all

1. In studying the investment on Research & Development of Thailand's pharmaceutical industry, a researcher found that the investment on R&D of each firm depended on firm's sale and profit. He obtained the following regression results:

$$RD_i = 0.472 + 0.621 \ln Sale_i + 0.095 \ln profit_i$$
$$se \quad (1.369) \quad (0.216) \quad (0.046)$$
$$R^2 = 0.09, n = 60$$

where RD = Investment on Research & Development, percentage of sales
Sales = Firm's sale, million dollars
Profit = Firm's profit, percentage of sales

1.1) Interpret each coefficients from the above equation.

1.2) Test the hypothesis whether an increase in firm's sale significantly increases investment on R&D with 0.05 and 0.01 level of significance.

1.3) Test the hypothesis whether firm's sale and profit jointly determine investment on R&D at 0.01 level of significance.

1.4) Do you think the preceding regression model is appropriate? Why?
Provide a suggestion on how the researcher can improve his model.

2. Consider the *demand function* of following models:

Model (1) $Y_i = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 X_{2i} + \alpha_3 X_{3i} + u_i$

Model (2) $Y_i = \alpha_0^{'} + \alpha_1^{'} X_{1i} + u_i^{'}$

Model (3) $\ln Y_i = \beta_0 + \beta_1 \ln X_{1i} + \beta_2 \ln X_{2i} + \beta_3 \ln X_{3i} + u_i^{"}$

where Y = Quantity demanded for good A, units
X_1 = Price of good A, Baht
X_2 = Consumer's income, Baht
X_3 = Price of good B, Baht

2.1) Suppose you want to do hypothesis testing for each estimator(α_i) in model (1) and (2). Provide appropriate null hypothesis (H_0) and alternative hypothesis (H_a) for every estimator based on demand theory. [ex. H_0: $\alpha_1 = 0$, H_0: $\alpha_1 > 0$, etc.]
Also, provide economic reasons for your hypothesis.

2.2) How can you test whether you should add/drop any variable to/from the model? Explain your method by comparing model (1) and (2).

2.3) Consider estimation results of model (1) and (3), explain the meaning. Are they different? How?

2.4) From model (3), how would you interpret the economic meaning of the coefficient of each independent variable? Explain.

2.5) From the preceding three models, is good B a substitute or complementary product to good A? How do you know? (Specify all possible cases)

3. For industrial sector in one country during 1980-2010, the production function (Y) is estimated as a function of capital (K) and labor (L). The following regression results are obtained.

Model (1)
$$\ln Y_t = 18.27 + 0.536 \ln L_t + 0.024 \ln K_t$$
$$R^2 = 0.9389, RSS = 0.0124$$

Model (2)
$$\ln(Y/L)_t = 2.13 + 1.12 \ln(K/L)_t$$
$$R^2 = 0.8087, RSS = 0.0153$$

3.1) Interpret the meaning of slope coefficients in model (1) and (2)

3.2) Test the hypothesis whether production function in that country is constant return to scale at 5% level of significance.

3.3) Can we compare R^2 value between the above two models? Why or why not?

4. Consumer price index (CPI) and Money supply (MS: $) of France during 1981-97 can be expressed by the following regression results:

Model (1):

$$CPI_t = 38.9691 + 0.2609\ MS_t$$

P-value (0.000) (0.001)

$R^2 = 0.9423, F = 245.08, n = 17$

Model (2):

$$\ln CPI_t = 1.4041 + 0.5889\ \ln MS_t$$

P-value (0.007) (0.000)

$R^2 = 0.9642, F = 403.60, n = 17$

Model (3):

$$\ln CPI_t = 3.9315 + 0.0028\ MS_t$$

P-value (0.002) (0.000)

$R^2 = 0.9284, F = 194.59, n = 17$

Model (4):

$$CPI_t = -192.9661 + 54.2126\ \ln MS_t$$

P-value (0.000) (0.003)

$R^2 = 0.9543, F = 313.41, n = 17$

4.1) From model (1)-(4), is money supply statistically significant at 95% confidence?

4.2) What is the rate of change and elasticity of CPI with respect to MS of each model? [The elasticity is computed at the average point of CPI and MS, $\overline{Y} = 92$ and $\overline{X} = 205$]

4.3) Can we use R^2 as a criteria to choose the most appropriate model among those four models? Why or why not?

5. The model for aggregate saving for one economy is offered as following:

$$S_t = \beta_1 + \beta_2 Y_t + \beta_3 P_t + \beta_4 I_t + u_t$$

where

S_t is aggregate saving for the economy at year t (billions baht)

Y_t is national income at year t (billions baht)

P_t is number of population at year t (million people)

I_t is interest rate at year t (%)

Its estimation is as following

Source	SS	df	MS	Number of obs =	39
				F(3, 35) =	419.62
Model	45984.1561	3	15328.052	Prob > F =	0.0000
Residual	1278.47932	35	36.5279806	R-squared =	0.9729
				Adj R-squared =	0.9706
Total	47262.6354	38	1243.75356	Root MSE =	6.0438

s	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
y	.4686708	.2203497	2.13	0.041	.0213371	.9160045
p	-.1202995	1.663044	-0.07	0.943	-3.496458	3.255859
i	509.4828	539.7863	0.94	0.352	-586.3417	1605.307
_cons	11.50964	8.120521	1.42	0.165	-4.97589	27.99518

And its correlation follows

	s	y	p	i
s	1.0000			
y	0.9852	1.0000		
p	0.9842	0.9947	1.0000	
i	0.9367	0.9329	0.9601	1.0000

5.1) From tables above, does model face any violation of assumption of CLRM? If so, what indicators do you use to observe that occurrence? Explain and discuss about the results to estimators and hypothesis testing after such violation.

5.2) What will you do to solve such problems? Illustrate your solutions with applicable measure.

5.3) If you still want to model this aggregate saving for the economy, which variables should you retain in the model given these information?

For question 6 - 8, you have to submit STATA results by using data in http://goo.gl/Jups4q

6. Use data from file **401ksubs.xls**

where nettfa = Net total financial asset, thousand dollars
age = Age of respondents, years
inc = Annual income of a family, thousand dollars
fsize = Family size, members

For the following questions, use only data from the families that have 1 member [command: keep if fsize==1]

6.1) How many families that have only 1 member? Compute basic statistics (e.g. mean, standard deviation, etc.) for nettfta, age, and inc.

6.2) Estimate model as following:

$$nettfa = \beta_1 + \beta_2 inc + \beta_3 age + u$$

Interpret the results and test hypothesis for each coefficient whether it is individually significant.

6.3) Test overall significance of the equation in 6.2)

6.4) Estimate model as following:

$$nettfa = \beta_1 + \beta_2 inc + \beta_3 age + \beta_4 inc^2 + \beta_5 age^2 + u$$

Test the hypothesis that H_0: $\beta_4 = \beta_5 = 0$ and H_a : either $\beta_4$or β_5 is not equal to zero

7. Consumption expenditure behavior in Thailand during the 1st quarter of 1993 to 4th quarter of 2014 can be expressed by:

$$C_t = \beta_1 + \beta_2 Y_t + u_t$$

where C_t = Aggregate consumption expenditure (Million Baht, at time t)
Y_t = Aggregare income (Million Baht, at time t)

7.1) Basing on economic theory, how can you interpret the coefficient in the given model and what is the possible interval for the value of β_2. Provide appropriate reasons.

7.2) Estimate the preceding regression, using data from file **NESDB_Q_CY**.
Predict the value of dependent variable and plot a graph showing the relationship between the predicted dependent variable and independent variable.

7.3) Suppose we believe that expenditure behavior in Thailand during 1993-1999 was different from that during 2000-2014 due to economic crisis. Plot a graph showing the relationship of Y and C in the distinct two periods. Test whether our statement is true by performing Chow Test at 0.05 level of significance.

8. To study the relationship between education's compensation and parents' education, we use the following model:

$$\ln(wage) = \beta_1 + \beta_2 educ + \beta_3 educ \times pareduc + \beta_4 exper + u$$

where wage = Labor's wage per hour, $
educ = Labor's education, years
pareduc = Parents' education [gen pareduc = fatheduc+motheduc], years
exper = Experience, years

8.1) What is the effect on *an increase in education* to *wage* (ln(wage))?

8.2) Estimate the preceding regression, using data from **htv.dta**. [you have to generate new variable, $educ \times pareduc$ (command: gen educpareduc = educ*pareduc)]

What is the difference in wage between those with 32-year parents' education and those with 24-year parents' education? Is this factor, parent education, significant?

8.3) Suppose we add one more variable, $exper^2$, to the preceding model and specify a new model:

$$\ln(wage) = \beta_1 + \beta_2 educ + \beta_3 educ \times pareduc + \beta_4 exper + \beta_5 exper^2 + u$$

Estimate above model and explain the effect of an increase in *experience* to *wage.*

8.4) Test the hypothesis that experience does not affect wage (at 5% level of significance).