EE325 Introductory Econometrics (Section 1 semester 1/2020)

Assignment 4

Instruction: Write your answer in either paper or digital paper. However, if you write on paper, please scan it and save as a PDF file. Submission is via BE-Moodle as a PDF file for both cases. (Please keep the file below 10MB as that is the maximum per file capacity for student.)

Due date: Friday, November 6,2020 (Before 10 P.M.)

1. From the data for 46 states in the United States for 1992, results of the regression are displayed as follows.

$$\begin{aligned} \ln C_i &= 4.30 - 1.34 \ln P_i + 0.17 \ln Y_i \\ se &= (0.91) \quad (0.32) \qquad (0.20) \\ \bar{R}^2 &= 0.27 \end{aligned}$$

where C_i = cigarette consumption, packs per year

P_i = real price per pack, \$ per pack

Y_i = real disposable income per capita, \$ per week

1.1) Do the estimation results follow the law of demand?

1.2) What is the elasticity of demand for cigarettes with respect to price? Is it statistically significant? If so, is it statistically different from 1?

1.3) What is the income elasticity of demand for cigarettes? Is it statistically significant? If not, what might be the reasons for it?

2. From estimating the regression equation on net financial wealth (nettfa), age of the survey respondent (age), and annual family income (inc) for people in the United States. The wealth and income variables are both recorded in thousands of dollars. The OLS estimation results for the model are given by

$$nettfa_i = \beta_1 + \beta_2 inc_i + \beta_3 age_i + u_i$$

```
reg nettfa inc age

      Source |       SS           df       MS      Number of obs   =     9,275
-------------+----------------------------------   F(2, 9272)      =    943.21
       Model |   6414618.8         2   3207309.4   Prob > F        =    0.0000
    Residual |  31528770.7     9,272  3400.42825   R-squared       =    0.1691
-------------+----------------------------------   Adj R-squared   =    0.1689
       Total |  37943389.5     9,274   4091.3726   Root MSE        =    58.313

------------------------------------------------------------------------------
      nettfa |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inc |   .9533566   .0252775    37.72   0.000     .9038072    1.002906
         age |   1.030777   .0591226    17.43   0.000     .9148838     1.14667
       _cons |  -60.69654   2.596333   -23.38   0.000    -65.78592   -55.60715
------------------------------------------------------------------------------
```

```
reg nettfa inc age agesq

      Source |       SS           df       MS      Number of obs   =     9,275
-------------+----------------------------------   F(3, 9271)      =    646.80
       Model |  6567017.15         3  2189005.72   Prob > F        =    0.0000
    Residual |  31376372.3     9,271  3384.35685   R-squared       =    0.1731
-------------+----------------------------------   Adj R-squared   =    0.1728
       Total |  37943389.5     9,274   4091.3726   Root MSE        =    58.175

------------------------------------------------------------------------------
      nettfa |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inc |   .9782522   .0254891    38.38   0.000      .928288    1.028216
         age |  -2.231489   .4897118    -4.56   0.000    -3.191432   -1.271547
       agesq |   .0377221   .0056214     6.71   0.000      .026703    .0487413
       _cons |   4.680388   10.08099     0.46   0.642    -15.08056    24.44134
------------------------------------------------------------------------------
```

2.1) Test the coefficient, in the first model, $\beta_3 < 1$ in the first model or not?

2.2) Due to estimation result by adding the age^2 variable or agesq. Perform the test whether we should include the quadratic term of the age variable or not? (Test for both t-test and F-test.) Also, interpret the meaning of this coefficient.

3. You are conducting an empirical investigation into the median prices of houses in 506 communities of a large metropolitan area. The sample data consist of 506 observations on the following observable variables:

P_i: the median house price in community i, in dollars;

NOX_i: the level of nitrous oxide in the air of community i, in parts per 100 million;

$DIST_i$: the weighted distance of community i from municipal area, in miles;

$ROOM_i$: the average number of rooms per house in community i;

$STRAT_i$: the average student-teacher ratio of schools in community i.

Researcher estimates the following model of median house price. The OLS estimation results for the model are given by

$$\begin{aligned} ln(P_i) &= 11.08 - 0.9535\, ln(NOX_i) - 0.1343\, ln(DIST_i) + 0.2545 ROOM_i - 0.05245 STRAT_i \\ se &= (0.3181)\ (0.1167) \qquad (0.04310) \qquad (0.01853) \qquad (0.005897) \\ RSS &= 35.1835 \quad TSS = 84.5822 \end{aligned}$$

3.1) Interpret each of the coefficient estimates in regression equation.

3.2) Test the individual significance of each of the slope coefficient estimates for $ln(NOX_i)$ and $ROOM_i$.

3.3) Find the R-squared, adjusted R-squared, and test the joint significance of all the slope coefficient estimates.

3.4) If researcher would like to test the proposition that the marginal effect of $ln(NOX_i)$ on $ln(P_i)$ equals the marginal effect of $ln(DIST_i)$ on $ln(P_i)$, write the restricted model and perform the test comparing restricted and unrestricted model, given that OLS estimation of this restricted regression equation yields a Residual Sum of Squares value = 41.9532.

4. Production function (Y) of the industrial sector in Thailand. It depends on the capital factor (K) and labor factor (L) in the years 1980-2010 with the following estimation.

<u>Model 1</u>:

$$ln\, Y_t = 18.27 + 0.536\, ln\, L_t + 0.024\, ln\, K_t$$

$$R^2 = 0.9389, RSS = 0.0124$$

<u>Model 2</u>:

$$ln \left(\frac{Y}{L}\right)_t = 2.13 + 1.12\, ln \left(\frac{K}{L}\right)_t$$

$$R^2 = 0.8087, RSS = 0.0153$$

4.1) Interpret the coefficients of the independent variables in models 1 and 2.

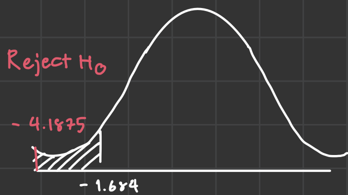
4.2) Test the hypothesis. Is the industrial production function characterized by constant return to scale? (Hint: you can perform any type of test that you see fit.)

4.3) Can we compare the R^2 value between the two regression models? Why?

1.1) $H_0 : \beta_2 \geq 0$

$H_1 : \beta_2 < 0$

$\alpha = 0.05$



$$t\ cal = \frac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}} = \frac{-1.34 - 0}{0.32} = -4.1875$$

$d.f. = 46 - 3 = 43$

∴ There is enough evidence to say that this result follow the law of demand $\alpha = 5\%$

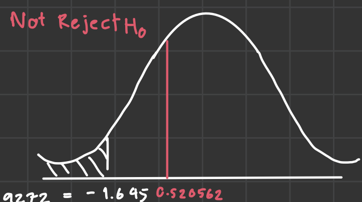
1.2) $\varepsilon_p = \frac{\%\Delta c}{\%\Delta p} = -1.34$

1.3) $C_I = \frac{\%\Delta c}{\%\Delta p} = 0.17$

2.1) $H_0 : \beta_3 \geq 1$

$H_1 : \beta_3 < 1$

$\alpha = 0.05$



$$t\ cal = \frac{\hat{\beta}_3 - \beta_3}{\sigma_{\hat{\beta}_3}} = \frac{1.0307 - 1}{0.0591} = 0.520562$$

$-t_{0.05,\,9272} = -1.645$

$df = 9275 - 3 = 9272$

∴ There is enough evidence to say that this result $\beta_3 < 1$ in the first model

2.2) F-test $\quad nettfe_i = \beta_1 + \beta_2 \ln c_i + \beta_3 age_i + \beta_4 age_i^2 + u_i$

$\beta_4 \neq 0 \quad (age^2)$

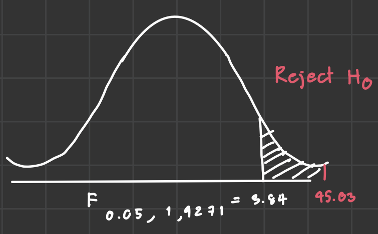
$nettfe_i = \beta_1 + \beta_2 \ln c_i + \beta_3 age_i + u_i$

$H_0 : \beta_4 = 0$

$H_1 : \beta_4 \neq 0$

$\alpha = 0.05, \quad m = k - r = 4 - 3 = 1$

$$F\ cal = \frac{RSS_R - RSS_{UR}/m}{RSS_{UR}/n-k} = \frac{R^2_{UR} - R^2_R / m}{(1 - R^2_{UR})\, n-k} = \frac{(31528770.7 - 31376372.3)/1}{31376372/9275-4} = 45.03$$



∴ There is enough evidence to say that the addition of age^2 to the model has significant at $\alpha = 5\%$

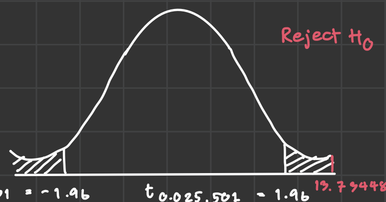
3.1) $\overset{\%}{\ln}\hat{}(P_i) = \overset{\hat{\beta}_1}{11.08} - \overset{\hat{\beta}_2}{0.9535}\,\overset{\%}{\ln}(NOX_i) - \overset{\hat{\beta}_3}{0.1343}\,\overset{\%}{\ln}(DIST_i) + \overset{\hat{\beta}_4}{0.2545}\,ROOM_i - \overset{\hat{\beta}_5}{0.05245}\,STRAT_i$

- $\hat{\beta}_1$ If $\ln(NOX_i) = 0$, $\ln(DIST) = 0$, $ROOM_i = 0$ and $STRAT_i = 0$, on average the median house price would be equal to $\$ e^{11.08} = \$ 64,860.88341$
- $\hat{\beta}_2$ As the level of nitrus oxide increase by 1%, on average the median house price decrease by amount of 0.9535%, other variable constant
- $\hat{\beta}_3$ As the weighted distance from municipal area increase by 1%, on average the median house price decrease by amount of 0.1343%, other variable constant
- $\hat{\beta}_4$ As the average number of rooms per house increase by 1 room per house, on average the median house price increase by amount of $0.2545 \times 100 = 25.45\%$, other variable constant
- $\hat{\beta}_5$ As the average student - teacher ration of schools increase by 1 student per teacher, on average the median house price decrease by amount of $0.05245 \times 100 = 5.245\%$.

3.2) $H_0 : \beta_4 = 0$ $\quad t_{cal} = \dfrac{\hat{\beta}_4 - \beta_4}{\sigma \hat{\beta}_4} = \dfrac{0.2545 - 0}{0.01853} = 13.73448$

$H_1 : \beta_4 \neq 0$

$\alpha = 0.05$ $\quad df = n - k = 506 - 5 = 501$

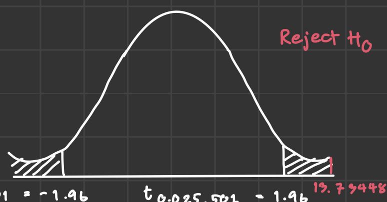


$\therefore$ There is enough evidence to say that $\ln(NOX)$ has significant on $\ln P_i$ at $\alpha = 0.05$

$H_0 : \beta_5 = 0$ $\quad t_{cal} = \dfrac{\hat{\beta}_5 - \beta_5}{\sigma \hat{\beta}_5} = \dfrac{-0.05245 - 0}{0.005897} = 13.73448$

$H_1 : \beta_5 \neq 0$

$\alpha = 0.05$ $\quad df = n - k = 506 - 5 = 501$



$\therefore$ There is enough evidence to say that $ROOM_i$ has significant on $\ln P_i$ at $\alpha = 0.05$

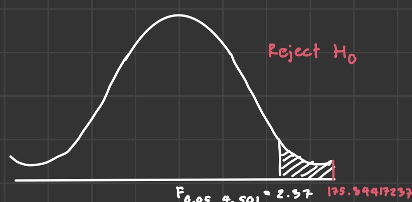
3.3) $R^2 = 1 - \dfrac{TSS}{RSS} = 1 - \dfrac{35.1835}{84.5822} = 1 - 0.416 = 0.584$

Adjust $R^2 = 1 - (1 - R^2)\dfrac{n-1}{n-k} = 1 - (1 - 0.341)\dfrac{(505)}{(501)} = 0.3358$

Test the joint ; $H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$

H_1 : Not all slope coefficients are simultaneously zero

$F_{cal} = \dfrac{ESS/k-1}{RSS/n-k} = \dfrac{R^2/k-1}{(1-R^2)n-k} = \dfrac{49.3957/4}{35.1835/501} = 175.8441$



$\therefore$ There is enough evidence to say that at least one parameter is not equal to zero, at $\alpha = 5\%$.

3.4) $H_0: \beta_2 = \beta_3$ or $\beta_2 - \beta_3 = 0$

$H_1: \beta_2 \neq \beta_3$ or $\beta_2 - \beta_3 \neq 0$

$\alpha = 0.05 \quad K = 5 \quad R = 4$

$RSS_R = 41.9532$

$$\ln(P_i) = \beta_1 + \beta_2 \ln(NOX_i) + \beta_3 \ln(DIST_i) + \beta_4(ROOM_i) + \beta_5(STRAT_i) + u_i$$

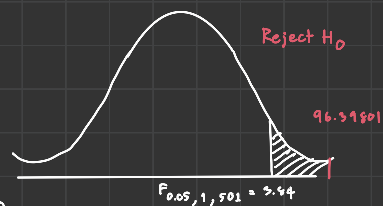
$$\ln(P_i) = \beta_1 + \beta_2 \ln(NOX_i) + \beta_2 \ln(DIST_i) + \beta_4(ROOM_i) + \beta_5(STRAT_i) + u_i$$

$$\ln(P_i) = \beta_1 + \beta_2 \left[\ln(NOX_i) + \ln(DIST_i)\right] + \beta_4(ROOM_i) + \beta_5(STRAT_i) + u_i$$

$$\ln(P_i) = \beta_1 + \beta_2 \ln(NOX_i \cdot DIST_i) + \beta_4(ROOM_i) + \beta_5(STRAT_i) + u_i$$

$$F_{cal} = \frac{RSS_R - RSS_{UR}/M}{RSS_{UR}/N-K} = \frac{R^2_{UR} - R^2_{R}/M}{(1-R^2_{UR})/n-k} = \frac{ESS_{UR} - ESS_{R}/M}{RSS_{UR}/n-K}$$

$$= \frac{41.9532 - 35.1835/1}{35.1835/506-5} = 96.39801$$



∴ There is enough evidence to say that the marginal effect of $\ln(NOX_i)$ on $\ln(P_i)$ equal the marginal effect of $\ln(P_i)$ at $\alpha = 5\%$.

4.1) Model 1 : $\hat{\ln y_t} = \underset{\hat\beta_1}{18.27} + \underset{\hat\beta_2}{0.531} \ln L_t + \underset{\hat\beta_3}{0.1024} \ln k_t$

- $\hat\beta_1$ If $\ln L_t = 0$ and $\ln k_t = 0$, on average output would be equal to $e^{18.27}$ units
- $\hat\beta_2$ As labor increase by 1%, on average output increase by amount of 0.536%
- $\hat\beta_3$ As capital increase by 1%, on average output increase by amount of 0.1024%.

Model 2 : $\ln\left(\frac{Y}{L}\right)_t = \underset{\hat\alpha_1}{2.13} + \underset{\hat\alpha_2}{1.12} \ln\left(\frac{K}{L}\right)_t$

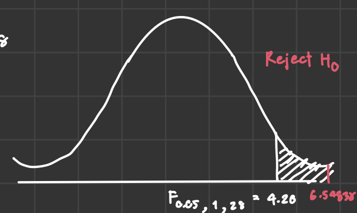
- $\hat\alpha_1$ If in $(K/L) = 0$, on average output per labour ratio would be equal to $e^{2.13}$ unit
- $\hat\alpha_2$ As capital per ratio increase by 1%, on average labour productivity increase by amount of 1.12%

4.2) $H_0: \beta_2 + \beta_3 = 1$

$H_1: \beta_2 + \beta_3 \neq 1$

$\alpha = 0.05 \quad M = k - r = 3 - 2 = 1$

$$F_{cal} = \frac{RSS_R - RSS_{UR}/m}{RSS_{UR}/n-k} = \frac{0.0153 - 0.0124/1}{0.0124/31-3} = 6.54838$$



∴ There is enough evidence to say that this industrial production function characterized by CRTS at $\alpha = 5\%$

4.3) Revise $var(ax \pm bY) = a^2 var(X) + b^2 var(Y) \pm 2ab\ cov(X_1, Y_1)$

$\sigma^2 = var$

$$\sigma(\hat\beta_2 + \hat\beta_3) = \sqrt{var(\hat\beta_2) + var(\hat\beta_3)}$$

t-test

$H_0: \beta_2 + \beta_3 = 1$

$H_1: \beta_2 + \beta_3 \neq 1$

$\alpha = 0.05$

$$t_{cal} = \frac{(\hat\beta_2 + \hat\beta_3) - (\beta_2 + \beta_3)}{\sigma(\hat\beta_2 + \hat\beta_3)} = \frac{(\hat\beta_2 + \hat\beta_3) - 1}{\sqrt{var(\hat\beta_2) + var(\hat\beta_3) + 2\,cov(\hat\beta_2 + \hat\beta_3)}} = \frac{(0.536 + 0.024) - 1}{\sqrt{var(\hat\beta_2) + var(\hat\beta_3) + 2\,cov(\hat\beta_2 + \hat\beta_3)}}$$