EE 325
Practice V Hints for answers

Students must submit question 1-2 to receive HW credits. Due date: Dec 9, 2019 by email. (Please convert to pdf file before you submit)

1. From the annual data from the U.S. manufacturing sector for 1899-1922, Dougherty obtained the following regression results:

$$\widehat{\log Y_t} = 2.81 - 0.53 \log K + 0.91 \log L + 0.047t$$
$$\text{se} = (1.38) \quad (0.34) \quad (0.14) \quad (0.021)$$
$$R^2 = 0.97 \quad F = 189.8 \tag{1}$$

Where Y = index of real output
K = index of real capital input
L = index of real labor input
t= time or trend

Using the same data, he also obtained the following regression:

$$\widehat{\log\left(\frac{Y}{L}\right)} = -0.11 + 0.11 \log\left(\frac{K}{L}\right) + 0.006t$$
$$\text{se} = \quad (0.03) \quad (0.15) \quad (0.006)$$
$$R^2 = 0.65 \quad F = 19.5 \tag{2}$$

a. Is there multicollinearity in regression (1)? How do you know?
Given the relatively high R2 of 0.97, the significant F value and the (economically speaking) improperly signed insignificant coefficient of log K, it may be that there is collinearity in the model.

b. How would you justify the functional form of regression (1)? (Hint: Cobb-Douglas production function)
It is a Cobb-Douglas type production function, as the given model can be written as:
$$Y = \beta_1 K^{\beta_2} L^{\beta_3} e^{\beta_4 t}$$

c. Interpret regression (1).
On average, over the sample period, a 1% increase in the index of the real labor input resulted in about 0.91% increase in the index of real output. The t variable in the model represents time. Very often, time is taken as a proxy for technical change. The coefficient of 0.47 suggests that over the sample period, on average, the rate of growth of real output (as measured by the output index) was about 4.7%.

d. What is the logic behind estimating regression (2)?
This equation implicitly assumes that there are constant returns to scale, that is, ($\beta_2 + \beta_3$) = 1. An incidental advantage of the transformation may be to reduce the collinearity problem.

e. If there was multicollinearity in regression (1), has that been reduced by regression (2)? How do you know?

Given that the capital-labor ratio coefficient is statistically insignificant, it appears that the collinearity problem has not been resolved.

f. If regression (2) is a restricted version of regression (1), what restriction is imposed by the author? (Hint: returns to scale.) How do you know if this restriction is valid? What test do you use? Show all your calculations.

As mentioned in (*d*), the author is trying to find out if there are constant returns to scale. One could use the *F* test to find out if the restriction is valid. But since the dependent variables in the two models are different, we cannot use the R^2 version of the *F* test. We need the restricted and unrestricted residual sums of squares to use the *F* test.

g. Are the R^2 values of the two regressions comparable? Why or why not? How would you make them comparable, if they are not comparable in the present form?
As noted in (f) the two R2's are not comparable. One could follow the procedure discussed in Multiple Regression chapter to render the two R2 values comparable.

2. For pedagogic purposes Hanushek and Jackson estimate the following model:

$$C_t = \beta_1 + \beta_2 GNP_t + \beta_3 D_t + u_t \qquad (1)$$

Where $C_t =$ aggregate private consumption expenditure in year t
$GNP_t =$ GNP in year t
$D_t =$ national defense expenditures in year t
The objective of the analysis being to study the effect of defense expenditure on other expenditures in the economy.

Postulating that $\sigma_t^2 = \sigma^2 (GNP_t)^2$, they transform (1) and estimate

$$C_t/GNP_t = \beta_1 (1/GNP_t) + \beta_2 + \beta_3 (D_t/GNP_t) + (u_t/GNP_t) \qquad (2)$$

The empirical results based on the data for 1946-1975 were as follows (standard errors in the parentheses):

$$\begin{aligned} \widehat{C}_t &= \underset{(2.73)}{26.19} + \underset{(0.0060)}{0.6248}\, GNP_t - \underset{(0.0736)}{0.4398} D_t \\ R^2 &= 0.999 \end{aligned}$$

$$\begin{aligned} \widehat{C_t/GNP_t} &= \underset{(2.22)}{25.92} (1/GNP_t) + \underset{(0.0068)}{0.6246} - \underset{(0.0597)}{0.4315} (D_t/GNP_t) \\ R^2 &= 0.875 \end{aligned}$$

a. What assumption is made by the authors about the nature of heteroscedasticity? Can you justify it?

The assumption made is that the error variance is proportional to the square of GNP, as is described in the postulation. The authors make this assumption by looking at the data over time and observing this relationship.

b. Compare the results of the two regressions. Has the transformation of the original model improved the results, that is, reduced the estimated standard errors? Why or why not?

The results are essentially the same, although the standard errors for two of the coefficients are lower in the second model; this may be taken as empirical justification of the transformation for heteroscedasticity.

c. Can you compare the two R^2 values? Why or why not?

No. The R^2 terms may not be directly compared, as the dependent variables in the two models are not the same.