EE325 HW 4 Answers

8.14

a) A priori, salary and each of the explanatory variables are expected to be positively related, which they are. The partial coefficient of 0.280 means, ceteris paribus, the elasticity of CEO salary is a 0.28 percent.
The coefficient 0.0174 means, ceteris paribus, if the rate of return on equity goes up by 1 percentage point (Note: not by 1 percent), then the CEO salary goes up by about 1.07 %. Similarly, ceteris paribus, if return on the firm's stock goes up by 1 percentage point, the CEO salary goes up by about 0.024%.

b) Under the individual, or separate, null hypothesis that each true population coefficient is zero, you can obtain the t values by simply dividing each estimated coefficient by its standard error. These t values for the four coefficients shown in the model are, respectively, 13.5, 8, 4.25, and 0.44. Since the sample is large enough, you can see that the first three coefficients are individually highly statistically significant, whereas the last one is insignificant.

c) To test the overall significance, that is, all the slopes are equal to zero, use the F test

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} = \frac{0.283/3}{0.717/205} = 27.02$$

Under the null hypothesis, this F has the F distribution with 3 and 205 df in the numerator and denominator, respectively. The p value of obtaining such an F value is extremely small, leading to rejection of the null hypothesis.

d) Since the dependent variable is in logarithmic form and the roe and ros are in linear form, the coefficients of these variables give semi elasticities, that is, the growth rate in the dependent variable for an absolute (unit) change in the regressor.

8.21

a) The own-price elasticity is -1.274.
b) From the t test, we obtain:

$$t = \frac{1.274 - 0}{0.527} = 2.4174$$

The p value of obtaining such a t statistic under the null hypothesis is about 0.034, which is small. Hence, we reject the hypothesis that the true price elasticity is zero.

c) We obtain:

$$t = \frac{-1.274 - (-1)}{0.527} = -0.5199$$

Since this t value is not statistically significant, we do not reject the hypothesis that the true price elasticity is unity.

d) Both the signs are expected to be positive, although none of these variables is statistically significant.

e) Perhaps our sample size is too small to detect the statistical significance of carnation prices on the demand for roses or that of income on the demand for roses. Moreover, expenditure on roses may be such a small part of total income that one may not notice the impact of income on demand for roses.

9.2

a) As per economic theory, the coefficients of X_2, X_5 are expected to be positive and that of X_3, X_8, and X_9 are expected to be negative. The coefficient of X_4 could be positive or negative, depending on wife's age and the number of children. Perhaps an interactive dummy of age and children under 6 or between 6 and 13 might shed more light on the relationship between age and desired hours of work.

b) Holding all other factors constant, one would expect that desired hours of work would be higher than the (common) intercept value of 1,286 hours. This coefficient, however, has a negative sign. However, since it is not statistically significant, we can say little about the impact of X_6 on (average) Y. As for X_7, its coefficient is expected to be positive, which it is. Not only that, it is statistically significant, as the t value is quite high.

c) Perhaps, this is due to collinearity between age and education, as well as collinearity of these variables with number of children. Also, notice that the model does not include years of schooling completed by husband.

9.21

a) Since the dummy enters in the log form, and since the log of zero is undefined, by redefining the dummy as 1 and 10, we can obtain logs of these numbers.
b) The regression results are as follows (t values in parentheses):

$$\ln(Savings)_t = -0.1589 + 0.6695 \ln Income_t - 0.00029 \ln D_t$$
$$t = (-0.2074) \quad (6.2362) \quad (-0.00505)$$
$$R^2 = 0.8780$$

Since **the dummy coefficient is not statistically significant**, for all practical purposes the two intercept terms are the same. The interpretation of the intercept coefficient of –0.1589 is that it represents the value of log of savings when all the regressors take a value of zero. Taking the antilog of this value, we obtain the value of 0.8531 (billions of dollars).
It may be interesting to compare the preceding regression results with the following results, which allow for the interaction effect:

$$\ln(Savings)_t = -2.0048 + 0.9288 \ln Income_t + 2.3278 \ln D_t - 0.2985(\ln D_t * Income_t)$$
$$t = (-0.26528) \quad (8.7596) \quad (3.9696) \quad (-3.9820)$$
$$R^2 = 0.9291$$

Now you get an entirely different picture, for the differential intercept and slope dummies are both significant. **For the 1982-1995 period, the MPS (marginal propensity to save) is 0.6303, whereas for the earlier period it is 0.9288.** By the same token, **the intercept term for the first period is negative but it is positive for the second period.**

9.28

a) Model (9.5.4) is a linear model, whereas the present one is a log-lin model. Therefore, the slope coefficients of the regressor in this model are to be interpreted as semi-elasticities. Qualitatively, both models give similar results. **Since the regressand in the two models are different, we cannot compare the two R2's directly.**

b) As noted in the chapter, **if we take the antilog of the dummy coefficient of 3.6772, what we obtain is the median savings in the period 1970-1981, holding all other factors constant. Now antilog (3.6772) = 39.5355. Thus, if income were zero, the median savings in 1970-1981 would be about 40 billion dollars. Again, one should interpret this number with a grain of salt.**

   Now **if we take the antilog of (3.6772 + 1.3971) = 159.8602, this would be median savings in the period 1982-1995, holding income constant.**

c) Regressing log of Y (savings) on X (income), the estimated error variances in the two periods are: $\hat{\sigma}^2 = 0.0122\ (df = 10)$ and $\hat{\sigma}^2 = 0.0182\ (df = 12)$ Under the null hypothesis that the variances of the two populations are the same, we form

$$F = \frac{0.0182}{0.0122} = 1.4918$$

   For 12, and 10 df in the numerator and denominator, respectively, this value is not significant. Hence, we can conclude that the two error variances are the same. Note that in the original model discussed in the chapter we regressed Y (not ln Y) on X. So, if there was heteroscedasticity in the original model and not in the log-lin model, it suggests that the log transformation may be more appropriate.

10.3

a) Although the numerical values of the intercept and the slope coefficients of PGNP and FLR have changed, their signs have not. Also, these variables are still statistically significant. These changes are due to the addition of the TFR variable, suggesting that there may be some collinearity among the regressors.

b) Since the t value of the TFR coefficient is very significant (the p value is only .0032), it seems TFR belongs in the model. The positive sign of this coefficient also makes sense in that the larger the number of children born to a woman, the greater the chances of increased child mortality.

c) This is one of those “happy” occurrences where despite possible collinearity, the individual coefficients are still statistically significant.

10.28

Since the R2 values in all the auxiliary regressions are uniformly high, it seems the data suffer from the multicollinearity problem.

There are probably too many substitute good variables in the equation. One could use only the composite substitute good price, price of chicken and disposable income as regressors. This was already done in Problem 7.19.

Creating a relative price variable, say the price of beef divided by the price of pork, might alleviate the collinearity problem.