Econ 715

Problem Set 3

Short Answer Problems

4.6. In Section 4.5, we used as an example testing the rationality of assessments of housing prices. There, we used a log-log model in price and assess [see equation (4.47)]. Here, we use a level-level formulation.

(i) In the simple regression model

\[ \text{price} = \beta_0 + \beta_1 \text{assess} + u, \]

the assessment is rational if \( \beta_1 = 1 \) and \( \beta_0 = 0 \). The estimated equation is

\[ \hat{\text{price}} = -14.47 + 0.976 \text{assess} \]

\( (16.27) \ (0.049) \)

\( n = 88, \text{SSR} = 165,644.51, R^2 = .820. \)

First, test the hypothesis that \( H_0 : \beta_0 = 0 \) against the two-sided alternative. Then, test \( H_0 : \beta_1 = 1 \) against the two-sided alternative. What do you conclude?

(ii) To test the joint hypothesis that \( \beta_0 = 0 \) and \( \beta_1 = 1 \), we need the SSR in the restricted model. This amounts to computing \( \sum_{i=1}^{n} (\text{price}_i - \text{assess}_i)^2 \), where \( n = 88 \), since the residuals in the restricted model are just \( \text{price}_i - \text{assess}_i \). (No estimation is needed for the restricted model because both parameters are specified under \( H_0 \).) This turns out to yield SSR = 209,448.99. Carry out the F test for the joint hypothesis.

(iii) Now, test \( H_0 : \beta_2 = 0, \beta_3 = 0, \text{and } \beta_4 = 0 \) in the model

\[ \text{price} = \beta_0 + \beta_1 \text{assess} + \beta_2 \text{lotsize} + \beta_3 \text{sqrft} + \beta_4 \text{bdrms} + u. \]

The \( R^2 \)-squared from estimating this model using the same 88 houses is .829.

(iv) If the variance of price changes with assess, lotsize, sqrft or bdrms, what can you say about the F test from part (iii)?

4.8. Consider the multiple regression model with three independent variables, under the classical linear model assumptions MLR.1 through MLR.6:

\[ y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u. \]

You would like to test the null hypothesis \( H_0 : \beta_1 - 3\beta_2 = 1 \).

(i) Let \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \) denote the OLS estimators of \( \beta_1 \) and \( \beta_2 \). Find \( \text{Var}(\hat{\beta}_1 - 3\hat{\beta}_2) \) in terms of the variances of \( \beta_1 \) and \( \beta_2 \) and the covariance between them. What is the standard error of \( \hat{\beta}_1 - 3\hat{\beta}_2 \)?

(ii) Write the \( t \) statistic for testing \( H_0 : \hat{\beta}_1 - 3\hat{\beta}_2 = 1 \).

(iii) Define \( \theta_1 = \beta_1 - 3\beta_2 \) and \( \theta_2 = \beta_1 - 3\beta_2 \). Write a regression equation involving \( \beta_0, \theta_1, \beta_2 \) and \( \beta_3 \) that allows you to directly obtain \( \hat{\theta}_1 \) and its standard error.

4.10. Regression analysis can be used to test whether the market efficiently uses information in valuing stocks. For concreteness, let return be the total return from holding a firm’s stock over the four-year period from the end of 1990 to the end of 1994. The efficient markets hypothesis says that these returns should not
be systematically related to information known in 1990. If firm characteristics known at the beginning of the period help to predict stock returns, then we could use this information in choosing stocks. For 1990, let $dkr$ be a firm’s debt to capital ratio, let $eps$ denote the earnings per share, let $netinc$ denote net income, and let $salary$ denote total compensation for the CEO.

(i) Using the data in RETURN.RAW, the following equation was estimated:

$$
\hat{\text{return}} = -14.37 + 0.321 dkr + 0.043 eps - 0.051 netinc + 0.0035 salary
$$

$$
(6.89) (0.201) (0.078) (0.0047) (0.0022)
$$

$$
n = 142, R^2 = 0.0395
$$

Test whether the explanatory variables are jointly significant at the 5% level. Is any explanatory variable individually significant?

(ii) Now, re-estimate the model using the log from for $netinc$ and $salary$:

$$
\hat{\text{return}} = -36.30 + 0.327 dkr + 0.069 eps - 4.74 \log(\text{netinc}) + 7.24 \log(\text{salary})
$$

$$
(39.37) (0.203) (0.080) (3.39) (6.31)
$$

$$
n = 142, R^2 = 0.0330
$$

Do any of your conclusions from part (i) change?

(iii) In this sample, some firms have zero debt and others have negative earnings. Should we try to use $\log(dkr)$ or $\log(eps)$ in the model to see if these improve the fit? Explain.

(iv) Overall, is the evidence for predictability of stock returns strong or weak?

Computer Exercises

Attach your computation results with your answer sheets.

C.4.4. In Example 4.9, the restricted version of the model can be estimated using all 1,388 observations in the sample. Compute the $R$-squared from the regression of $bwght$ on $cigs$, parity, and $faminc$ using all observations. Compare this to the $R$-squared reported for the restricted model in Example 4.9.

C.4.6. Use the data in WAGE2.RAW fro this exercise.

(i) Consider the standard wage equation

$$
\log(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{exper} + \beta_3 \text{tenure} + u.
$$

State the null hypothesis that another year of general workforce experience has the same effect on $\log(\text{wage})$ as another year of tenure with the current employer.

(ii) Test the null hypothesis in part (i) against a two-sided alternative, at the 5% significance level, by constructing a 95% confidence interval. What do you conclude?

C.4.8. The data set 401KSUBS.RAW contains information on net financial wealth($nettfa$), age of the survey respondent($age$), annual family income($inc$), family size($fsize$), and participation in certain pension plans fro people in the United States. The wealth and income variables are both recorded in thousands of dollars. For this question, use only the data for single-person households (so $fsize = 1$).

(i) How many single-person households are there in the data set?
(ii) Use OLS to estimate the model

\[ \text{nettfa} = \beta_0 + \beta_1 \text{inc} + \beta_2 \text{age} + u. \]

and report the results using the usual format. Be sure to use only the single-person households in the sample. Interpret the slope coefficients. Are there any surprises in the slope estimates?

(iii) Does the intercept from the regression in part (ii) have an interesting meaning?

(iv) Find the $p$-value for the test $H_0 : \beta_2 = 1$ against $H_1 : \beta_2 < 1$. Do you reject $H_0$ at the 1% significance level?

(v) If you do a simple regression of nettfa on inc, is the estimated coefficient on inc much different from the estimate in part (ii)? Why or why not?