KAN-CEAPV2505U

Econometric Analysis of Firm Data

July 30, 2025

Copenhagen Business School
Department of Economics
Ralf A. Wilke
Mette Franck
2025/2026, Autumn semester

Problem Set 2: OLS - Topics

1. Which of the following are consequences of heteroskedasticity?
   (a) The OLS estimators, $\hat{\beta}_j$ are inconsistent.
   (b) The usual F statistic no longer has an F distribution.
   (c) The OLS estimators are no longer BLUE.
2. Use the data hprice1.dta and consider the house price equation

$$log(price) = \beta_0 + \beta_1 log(lotsize) + \beta_2 log(sqrft) + \beta_3 bdrms + u.$$

   (a) Estimate the model by OLS and compare the usual standard errors and heteroscedasticity robust standard errors.

(b) Apply the full White test for heteroscedasticity. Using the F test form of the statistic, obtain the p-value. What do you conclude?

3. Use the data in CEOSAL2.dta for this exercise.

   (a) Estimate the model

$$\begin{aligned} log(salary) &= \beta_0 + \beta_1 log(sales) + \beta_2 log(mktval) + \beta_3 profmarg \\ &\quad + \beta_4 ceoten + \beta_5 comten + u. \end{aligned}$$

   (b) When $ceoten^2$ and $comten^2$ are added to the model of part (a), the R-squared increases from 0.353 to 0.375. Is this evidence of functional form misspecification?

   (c) Apply the RESET test for functional form misspecification to the model of part (a).

4. Use the data in GPA2.dta for this exercise.

   (a) Estimate the model

$$sat = \beta_0 + \beta_1 hsize + \beta_2 hsize^2 + u,$$

   where sat is the score of a standardized test widely used for university admissions in the United States and $hsize$ is the size of the graduating class (in hundreds). Write the results in the usual form. Is the quadratic term significant?

   (b) Using the estimated equation from part (a), what is the "optimal" high school size? Justify your answer.

   (c) Find the estimated optimal high school size, using $\log(sat)$ as the dependent variable. Is it much different from what you obtained in part (b)?

   (d) Use $size = 100 * hsize$, the size of the graduating class, instead of $hsize$ to re-estimate the model in part (a). Confirm the effects on results after re-scaling the data.

These problems have been partly taken from the textbook "Introductory Econometrics" by J.Wooldridge, 7th edition, 2020.