FINM3123 Introduction to Econometrics

Chapter 8 Exercises

1. Consider a linear model to explain monthly beer consumption:

$$beer = \beta_0 + \beta_1 inc + \beta_2 price + \beta_3 educ + \beta_4 female + u$$

$\mathrm{E}(u|inc, price, educ, female) = 0$

$\mathrm{Var}(u|inc, price, educ, female) = \sigma^2 inc^2$.

Write the transformed equation that has a homoscedastic error term.

2. The variable *smokes* is a binary variable equal to one if a person smokes, and zero otherwise. Using the data in SMOKE.RData, we estimate a linear probability model for *smokes*:

$$\begin{aligned}\widehat{smokes} = \underset{\substack{(.855)\\ [.856]}}{.656} &- \underset{\substack{(.204)\\ [.207]}}{.069}\log(cigpric) + \underset{\substack{(.026)\\ [.026]}}{.012}\log(income) - \underset{\substack{(.006)\\ [.006]}}{.029}educ \\ &+ \underset{\substack{(.006)\\ [.005]}}{.020}age - \underset{\substack{(.00006)\\ [.00006]}}{.00026}age^2 - \underset{\substack{(.039)\\ [.038]}}{.101}restaurn - \underset{\substack{(.052)\\ [.050]}}{.026}white\end{aligned}$$

$$n = 807,\ R^2 = .062,$$

where

cigpric = the per-pack price of cigarettes (in cents)l

income = annual income.

educ = years of schooling.

age = measured in years.

restaurn = a binary indicator equal to unity if the person resides in a state with restaurant smoking restrictions.

white = a binary variable equal to one if the respondent is white, and zero otherwise.

Both the usual and heteroskedasticity-robust standard errors are reported.

i) Are there any important differences between the two sets of standard errors?

ii) Holding other factors fixed, if education increases by four years, what happens to the estimated probability of smoking?

iii) At what point does another year of age reduce the probability of smoking?

iv) Interpret the coefficient on the binary variable *restaurn*.

v) Person number 206 in the data set has the following characteristics: *cigpric* = 67.44, *income* = 6,500, *educ* = 16, *age* = 77, *restaurn* = 0, white = 0, and *smokes* = 0. Compute the predicted probability of smoking for this person and comment on the result.

3. Use VOTE1.RData for this exercise.
   i) Estimate a model with *voteA* as the dependent variable and *prtystrA*, *democA*, log(*expendA*), and log(*expendB*) as independent variables. Obtain the OLS residuals, $\widehat{u_i}$, and regress these on all of the independent variables. Explain why you obtain R^2 = 0.
   ii) Now, compute the Breusch-Pagan test for heteroskedasticity. Use the F statistic version and report the p-value.
   iii) Compute the special case of the White test for heteroskedasticity, again using the F statistic form. How strong is the evidence for heteroskedasticity now?

4. Use the data in MEAP00.RData to answer this question.
   i) Estimate the model

$$math4 = \beta_0 + \beta_1 lunch + \beta_2 \log(enroll) + \beta_3 \log(exppp) + u$$

   by OLS and obtain the usual standard errors and the fully robust standard errors. How do they generally compare?
   ii) Apply the special case of the White test for heteroskedasticity. What is the value of the F test? What do you conclude?
   iii) Obtain $\hat{g}_i$ as the fitted values from the regression $\log(\hat{u}_i^2)$ on $\widehat{math4}_i$, $\widehat{math4}_i^2$, where $\widehat{math4}_i$ are the OLS fitted values and the $\hat{u}_i$ are the OLS residuals. Let $\hat{h}_i = \exp(\hat{g}_i)$. Use the $\hat{h}_i$ to obtain WLS estimates. Are there big differences with the OLS coefficients?
   iv) Obtain the standard errors for WLS that allow misspecification of the variance function. Do these differ much from the usual WLS standard errors?
   v) For estimating the effect of spending on *math4*, does OLS or WLS appear to be more precise?