Práctica 2
GRETL. Simple Regression

Estadística-II. INTRODUCCIÓN a la ECONOMETRÍA. UC3M

1. Consider Wooldrige's dataset ATTEND.RAW to study the relation between the final grades for college students and class assistance and homework. Class assistance is measured as the rate (in percentage terms) of attended classes by the student ($atndrte$), homework is the rate of homework handed in by the student ($hwrte$) and final grades are standardized ($stndfnl$). Use the linear model

$$stndfnl = \beta_0 + \beta_1 atndrte + \beta_2 hwrte + \varepsilon \tag{1}$$

to explain the relation between the proportion of attended classes and the final grade, where β_0 and β_1 are unknown parameters and ε is an error term.

(a) Obtain and interpret the OLS estimates $\hat{\beta}_1$ y $\hat{\beta}_2$.

(b) Obtain $\widehat{Cov}(stndfnl, atndrte)$ and $\widehat{Var}(atndrte)$. Then, get the OLS for δ_1 in the "short" model,

$$stndfnl = \delta_0 + \delta_1 atndrte + \varepsilon',$$

(ε' is another error term) using

$$\hat{\delta}_1 = \frac{\widehat{Cov}(stndfnl, atndrte)}{\widehat{Var}(atndrte)}.$$

Use also gretl to compute $\hat{\delta}_1$.

(c) Compute $\widehat{Cov}(stndfnl, hwrte)$ and using a. y b., and the fact that:

$$\hat{\delta}_1 = \hat{\beta}_1 + \hat{\beta}_2 \frac{\hat{C}(atndrte, hwrte)}{\hat{V}(atndrte)},$$

obtain $\hat{\delta}_1$ and check that you get the same resultsas as in b.

- Coment on the consequences of omitting $hwrte$ in the short model over the interpretation of the effect of school attendance on the final grade.
- Show how to obtain in gretl directly the ratio:

$$\frac{\widehat{Cov}(atndrte, hwrte)}{\widehat{Var}(atndrte)},$$

(d) Compare standard errors for the estimators of the coefficients of $atndrte$ in both models, "short" and "long".

(e) Estimate, using OLS in gretl, the model,

$$atndrte = \gamma_0 + \gamma_1 hwrte + u,$$

and store the OLS residuals, $\hat{u}_i$.
Fit the model

$$stndfnl = \beta_0 + \beta_1 \hat{u} + v$$

and set $\tilde{\beta}_1$ as the adjusted coefficient for $\hat{u}$.
► Compare $\tilde{\beta}_1$ with the value of the coeffcient of $atndrte$ in (a), $\hat{\beta}_1$, and in (b), $\hat{\delta}_1$. Comnent the result. How can you interpret $\hat{u}$?

2. In this exersice, we use Wooldrige's dataset KIELMC.RAW to study the effect on the instalation of an incinirator on the prices of near-by houses (This dataset was originally used in K.A. Kiel and K.T. McClain (1995), "House Prices During Siting Decision Stages: The Case of an Incinerator from Rumor Through Operation," *Journal of Environmental Economics and Management* 28, 241-255). The set of variables is:

$price =$	house selling price
$age =$	house age
$area =$	size of house in square feet
$dist =$	distance in feet from house to incinirator
$baths =$	number of bathrooms in the house
$nbh =$	code for neighborhood quality from 1 to 6 (1: best)

and the model is

$$\log(price) = \beta_0 + \beta_1 dist + \beta_3 age + \beta_4 age^2 + \beta_5 \log(area) + \beta_6 baths + \varepsilon$$

(a) Estimate the model and interpret each of the coefficients.

(b) Compute an estimation for the expected elasticity, ceteris paribus, of the price with respect to the age of the house for house at the critical age (that is, the age at which the sign of the partial effect changes).

(c) Compare β_1 with an estimator of the slope in a "short" model which only uses *dist* as regressor. Which conclusions can be drawn from this analysis?