

Problem Set #3
due Thursday, Feb.28

1. The data set auto.dta contains data describing new car models sold in the US in 1978. Variables include price (in \$), weight (lbs), foreign (0= domestic, 1= foreign), mpg (miles per gallon). You can download it from <http://www.stata-press.com/data/r9/auto.dta> then read it into Stata by giving the command "use auto.dta" .

Create a variable which indicates if a car is domestically produced using the Stata command generate domestic = 1-foreign

Estimate the following three equations:

$$(1) \text{ price} = \beta_0^S + \beta_1^S \text{weight} + u^S \quad (S)$$

$$(2) \text{ price} = \beta_0^L + \beta_1^L \text{weight} + \beta_2^L \text{domestic} + u^L \quad (L)$$

$$(3) \text{ price} = \beta_0^{RL} + \beta_1^{RL} \text{weight} + \beta_2^{RL} \text{domestic} + \beta_3^{RL} \text{mpg} + u^{RL} \quad (RL)$$

a) What proportion of the variation in price is explained by the right hand side variables in equation (L)?

b) Test the hypothesis that β_2^{RL} is zero in equation (RL) against the alternative that β_2^{RL} is not zero at a level of significance α of .05.

c) Test the joint hypothesis that both β_1^{RL} , and β_2^{RL} are zero in equation (RL) against the alternative that they are not both zero at $\alpha = .05$. Report the test statistic.

d) Which are heavier on average, domestic or foreign cars (in this sample)? Show how you can use the omitted variable bias formula and the results of the estimated equations (S) and (L) to calculate the difference in weight between the average domestic car and the average foreign car.

e) A 1978 Volvo 260 weighed 3170 lbs., was foreign and got 17 miles per gallon in gas mileage. Use equation (RL) to predict how much your father would have had to pay for a new Volvo 260 in 1978.

f) What was the actual price of a Volvo 260 in 1978?

What is the sign of the estimated residual for this car from equation (RL)?

Was the Volvo 260 overpriced according to your calculations?

A fair number of Volvo 260s were sold. How can you explain their popularity?

g) An engineering student looks at your estimate for β_1 in equation (RL) and asks you if it means he can increase the value of a car by weighing it down with bags of wet sand in the rear trunk. Can he?

If not, how would you explain to the engineering student the interpretation of β_1 ?

2. (This has nothing to do with cars, necessarily)

Prove that the R^2 from a regression of Y on X in the sample is the same as the R^2 from a regression of Y on the predicted value, \hat{Y} , in the same sample.

You can always demonstrate this using some data in Stata. (After a regression the command "predict yhat, xb" will create a variable "yhat" of predicted values.) Of course a demonstration isn't a proof.

3. If all the observations in the sample lie exactly on a horizontal line (parallel to the X axis), what will the value of the R^2 be from a regression of Y on X ?

Prove it.