Homework Assignment

UNIVERSITY OF EXETER
BUSINESS SCHOOL

Maximum marks: 100

Applied Econometrics for Business (BEE3071)

Module convenor: Pradeep Kumar

Deadline: Nov 14, 2024 at 3pm

By the deadline you must have submitted a document (MS Word or pdf) containing your answers to the questions, including any tables and graphs. At the end of the document include the Stata commands you used to obtain the answers. Note: Only ONE document needs to be submitted, which will include both your answers and the Stata code.

Materials to be supplied: Data file.

Read this before starting: You may use all lecture slides, Stata help files, any notes (handwritten or printed), or any textbooks. Use of approved calculators is allowed. Whenever conducting a test, use a 5% significance level (95% confidence level) unless stated otherwise. **You may not confer with each other, or use external sources to answer the questions.**

Detailed Instructions:

Answer all questions.

Submission is online via ELE. The submitted document will contain your answers to the questions, including any output, tables and graphs. You must indicate which question your commands relate to. You may only submit once; it is not possible to alter your documents and re-submit at a later time.

The data corresponding to this homework is uploaded on the ELE page (*prod.dta*).

Marking will follow the university's assessment guidelines. Emphasis is placed on the correct interpretation of summary statistics, regression coefficients and hypothesis tests. For graphical analysis, emphasis is placed on the conclusions drawn from a graph, as well as a mention of the limitations, if any.

Unless mentioned otherwise, assume a significance level of 5% (i.e. 95% confidence level) for all statistical tests. Always state the null hypothesis and alternative hypothesis of the test, as well as your decision at the stated level of significance.

Introduction

You are given a data set (*prod.dta*) on 158 manufacturing firms which contains information on the cost of production, output quantity, prices and cost shares. The data contains following variables for each firm:

Variable	**Description**
id	Id of a firm
cost	Cost of production (in 10,000 pounds)
q	Quantity of output produced (number of units)
pl	Price of labour (in pounds)
pf	Price of fuel (in pounds)
pk	Price of capital (in 1,000 pounds)
sl	Share of the cost incurred by labour
sf	Share of the cost incurred by fuel
sk	Share of the cost incurred by capital
large	=1 if the firm is large, =0 if the firm is small

I Basic Descriptive Statistics [35 marks]

1. [6 marks] Report the mean, median and skewness of the cost of production (*cost*) for firms in the data. Is the skewness positive? Interpret it.

2. [8 marks] Calculate the fuel consumption in pounds by multiplying the total cost (*cost*) and the fuel share (*sf*). Store your result in a new variable, *fuelcost*. Using a graph of your choice, compare and discuss the fuel consumption (*fuelcost*) of small (*large=0*) and large (*large=1*) firms.

3. [5 marks] How many firms in the data are small (*large=0*) and spend more than the median cost share on both labour (*sl*) and fuel (*sf*)?

4. [10 marks] Create a new variable, *size*, with three categories: $q \leq 1961$; $1961 < q \leq 12542$; and $q > 12542$, where q is the quantity of output produced. Compute the average price of labour for each of the categories.

5. [6 marks] Is the labour share of cost (*sl*) same for large and small firms, on average? Interpret your result.

II Linear Regression Analysis [65 marks]

$$cost = \beta_0 + \beta_1 q + \beta_2 pk + \beta_3 pf + \beta_4 pl + error \tag{1}$$

1. [10 marks] Estimate a linear regression model with *cost* as the dependent variable and *q, pk, pf* and *pl* as the independent variables. Interpret the R^2 value. Also, interpret the meaning of the following coefficients in the regression model: β_0 and β_3.

2. [12 marks] Using the regression model (1), predict the cost incurred to produce 20,000 units of ourput (*q=20000*), where price of capital is 80,000 pounds (*pk=80*), price of fuel is 40 pounds (*pf=40*) and price of labour is 9,000 pounds (*pl=9000*). Create a 95% confidence interval around the predicted cost.

3. [20 marks] Estimate the regression model in (1) after adding two additional variables: q^2 and $q \times pk$. Using this new regression model answer the following:

   (a) [10 marks] What is the marginal effect of quantity of output (q) on the cost of production (*cost*) when *q=9000* and *pk=60*?

   (b) [5 marks] Interpret the coefficent on the variable $q \times pk$.

   (c) [5 marks] Compare and discuss the adjusted R^2 of this new regression model with the one in equation (1).

4. [15 marks] Estimate a new regression model with *log(cost)* as the dependent variable and *q, pk, pf* and *pl* as the independent variables. Use a scatter plot to compare the fitted values of cost using this new regression model to the actual values of cost in data. Repeat the same exercise using fitted values of cost from the regression model in equation (1) to do a scatter plot. Use the two scatter plots to compare which is a better model for predicting *cost*?

5. [8 marks] Is the relationship between cost of production (*cost*) and price of fuel (*pf*) causal? Discuss.

——— *End of assignment* ———