SOK-3011 Linear Models in Economics

Assignment 1

Instructions:

You will analyze a dataset containing firm-level production and cost data. Using this data, complete the following exercises, providing clear explanations and justifications for each step.

Total credits: 50 points. You are required to attempt all questions. Submit a comprehensive report that includes your answers to each question, along with all relevant estimations, interpretations, and discussions. Additionally, include the code/scripts used for your estimations.

Question 1: Data Collection (5 points)

Collect a firm-level dataset that includes at least the following variables: Output–The quantity of goods produced; Inputs–At least two types of inputs; Costs–Total production cost or variable cost, or the input prices. Data on the output price or revenue will also be useful. Ensure that the dataset contains a sufficient number of observations (no. of firms > 25) for meaningful analysis.

Exercise

- A brief description of the dataset (source, number of firms, number of variables, etc.).

- Descriptive statistics of the dataset (mean, standard deviation, min, max for each variable).

Important note: If you are unable to collect raw data, you may use datasets available in R packages, such as *sfaR*, *micEcon*, *rDEA*, *deaR*, *Benchmarking*, or others. A good resource to find relevant datasets is (https://vincentarelbundock.github.io/Rdatasets/datasets.html). Even if using an existing dataset, you must describe it properly, citing the original sources and explaining its purpose, industry, and time period.

Question 2: Estimation of Production Functions (30 points)

In this part, you estimate and analyze at least two (preferably more) of the following four production functions using your dataset: the Linear Production Function; Cobb-Douglas Production Function; Quadratic Production Function; Translog Production Function.

Exercise:

1. For each specification, estimate the production function using your data. Interpret the estimated coefficients. Analyze and discuss (i) productivity, (ii) returns to scale, and (iii) input substitution, based on the estimated model. Further, evaluate if the observed input choices align with cost-minimizing or profit-maximizing behavior.

2. Perform a goodness-of-fit analysis. Apply Ramsey's Regression Equation Specification Error Test (RESET) to check for model specification errors. Explain which estimated functional form performs better, using both statistical criteria and economic reasoning.

Question 3: Estimation of Cost Function or Profit function (10 points)

In this part, you may choose to work with either the cost function or the profit function, using a Cobb-Douglas specification.

Exercise:

1. Estimate the chosen cost or profit function using your dataset. Interpret the estimated coefficients and assess whether the model aligns with the theoretical economic properties of the cost or profit function. This includes examining the relationships with input and output prices, and applying Shephard's Lemma for cost functions or Hotelling's Lemma for profit functions, as appropriate.

2. Examine if the estimates provide evidence of optimal input choices, consistent with cost-minimizing or profit-maximizing behavior. Are these observations consistent with the results obtained from the primal approach in Question 2, in which you estimated the production functions?

Question 4: Discussion and recommendations (5 points)

Based on your analysis, write a short essay (approximately 500 words) discussing the implications of your findings for understanding firm behavior in the industry from which the data is drawn. Address any limitations of your analysis, including potential issues with data quality, estimation techniques, or model assumptions. Highlight any recommendation or policy implications based on your analysis (e.g., how firms in the industry could improve efficiency).