Chapter 7: Estimation of Cost and Production Functions

Quantitative Methods for Regulation and Competition

Motivation: Production and Cost Functions

- Evaluate efficiency:
  - Increasing returns, economies of scale
  - Complementarity/substitutability between inputs
  - Synergies and mergers
- Productivity analysis:
  - Measurement of productivity across firms or over time
  - Effects of policy (deregulation, tariffs,…)
  - Returns of R&D, returns to adoption of new technology

Today's Lecture

- Estimating production functions:
  - Returns to scale and technical change
  - Panel data analysis
- Estimating cost functions:
  - Comparative efficiency
- Index numbers and TFP

Model

- Suppose that production follows a Cobb-Douglas function:

$$Q_i = e^{\alpha} K_i^{\beta_k} L_i^{\beta_l} e^{u_i} \text{ where}$$

Q_i is the output of plant (or firm) i

L_i is the labour input (or more generally a variable input) of plant (or firm) i

K_i is the capital input (or more generally a fixed input) of plant (or firm) i

u_i is an error term

α, β_l, β_k are parameters to be estimated

Model (2)

- Remarks:
  - We might also include: material, energy, different types of labour
  - Error term includes: technology or management differences, measurement errors, variation in external factors

- Taking logs we obtain:

$$y_i = \alpha + \beta_k k_i + \beta_l l_i + u_i \text{ where}$$

y_i=ln (Q_i), l_i=ln (L_i) and k_i=ln (K_i)

Measurement Problems

- Production:
  - If multiproduct firm then output usually sales divided by a price index (more on this later)
  - Or using a quantity index (more on this later)
- Labour:
  - Usually hours or person per year
  - May adjust for types of labour
- Other non-capital inputs:
  - E.g. fuel (litres)
  - Other (expenditures divided by price index)
- Capital:
  - Investment at different times computing depreciation

Returns to Scale

- Estimating the previous model by OLS one finds estimates of α, β_l and β_k
- β_l and β_k are the elasticities of output wrt input
- Remember that with Cobb-Douglas model:

$$Q_i = AK^{\beta_k} L^{\beta_l} \text{ exhibits...}$$

$$\text{CRS iff } \beta_l + \beta_k = 1$$

$$\text{IRS iff } \beta_l + \beta_k > 1$$

$$\text{DRS iff } \beta_l + \beta_k < 1$$

- We can perform F-tests to determine the case

Technical Change Measurement

- With time series (or panel) data, we can estimate technical change:

$$y_t = \alpha + \beta_k k_t + \beta_l l_t + \beta_t t + u_t$$

where t is a time trend (t=1,2,…T)

- Interpretation:
  - β_t provides an estimate of the annual percentage change in output resulting from technical change (one year)

Violations of the OLS Assumptions

- What if the model is

$$y_i = \alpha + \beta_k k_i + \beta_l l_i + \beta_x X_i + u_i$$

but we estimate

$$y_i = \tilde{\alpha} + \tilde{\beta}_k k_i + \tilde{\beta}_l l_i + u_i$$

- The our estimates are biased if…
  (a) the omitted variable has an effect on the dependent variable and
  (b) the omitted variable is correlated with an included variable

- X may not be observable ("unobserved heterogeneity"):
  - Managerial efficiency:
    - Better managers may need less labour to produce same output
    - These firms produce more with less labour and OLS will underestimate β_l
  - Productivity shock:
    - A higher productivity shock may mean more labour
    - OLS will attribute all the increase in output to the change in labour and OLS will overestimate β_l

Panel Data

- Suppose that we have panel data:
  - Several firms followed at different points in time
  - Time-series of a cross section
- Assume that:

$$y_{i,t} = \alpha + \beta_k k_{i,t} + \beta_l l_{i,t} + \gamma_i + \varepsilon_{i,t} \text{ where}$$

$$\gamma_i \text{ is firm - specific and}$$

$$\varepsilon_{i,t} \text{ satisfies classical assumptions}$$

- We will look at two different models:
  - Fixed effects
  - Random effects

Panel Data: Fixed Effects

- Assume that γ_i is predetermined:
  - Not result of random variation, fixed over time
  - Can account for different managerial capacity for example
  - Might be correlated with labour though

- If this is true then panel data….
  - Allows us to obtain unbiased estimates for the variables included in the model
  - Correct for unobservable heterogeneity
  - Acts like a dummy shifting the intercept for each individual (impossible to do with cross section)

Fixed Effects: Methods

- Take first differences and OLS:

$$y_{i,t} = \alpha + \beta_k k_{i,t} + \beta_l l_{i,t} + \gamma_i + \varepsilon_{i,t}$$

$$y_{i,t-1} = \alpha + \beta_k k_{i,t-1} + \beta_l l_{i,t-1} + \gamma_i + \varepsilon_{i,t-1}$$

$$y_{i,t} - y_{i,t-1} = \beta_k (k_{i,t} - k_{i,t-1}) + \beta_l (l_{i,t} - l_{i,t-1}) + (\varepsilon_{i,t} - \varepsilon_{i,t-1})$$

- Compute deviation from means and OLS:

$$y_{i,t} = \alpha + \beta_k k_{i,t} + \beta_l l_{i,t} + \gamma_i + \varepsilon_{i,t}$$

$$\overline{y}_i = \alpha + \beta_k \overline{k}_i + \beta_l \overline{l}_i + \gamma_i + \overline{\varepsilon}_i$$

$$y_{i,t} - \overline{y}_i = \beta_k \left(k_{i,t} - \overline{k}_i\right) + \beta_l \left(l_{i,t} - \overline{l}_i\right) + \left(\varepsilon_{i,t} - \overline{\varepsilon}_i\right)$$

- Include a dummy variable for each firm (equivalent to compute deviations from means)

Fixed Effects: Extensions

- We may also include a time-variant variable:
  - Common effect impacting on all individuals
  - Can account for productivity shocks
- We can then use:

$$y_{i,t} = \alpha + \beta_k k_{i,t} + \beta_l l_{i,t} + \gamma_i + \gamma_t + \varepsilon_{i,t} \text{ where}$$

γ_i is firm-specific and γ_t is time-specific and

$\varepsilon_{i,t}$ satisfies classical assumptions

Panel Data: Random Effects

- Firm-specific component γ_i might be random:
  - Realisation of a specific random variable
  - Individual shifts of the intercepts are not deterministic
- We can estimate

$$y_{i,t} = \alpha + \beta_k k_{i,t} + \beta_l l_{i,t} + \gamma_i + \varepsilon_{i,t}$$

Where γ_i is also considered "error"

- Here, assume that:

$$\mathrm{var}(\gamma_i) = \sigma_\gamma, Cov(\gamma_i, \gamma_j) = 0 \text{ for all i} \neq \text{j}$$
$$Cov(\gamma_i, l_{j,t}) = 0, Cov(\gamma_i, k_{j,t}) = 0 \text{ for all i, j, t}$$

- Therefore OLS is unbiased but not efficient (because of serial correlation). Estimation using GLS method

Fixed Effects or Random Effects

- Fixed effects:
  - Estimate a lot of parameters
  - A lot of variation in the data is used to estimate these parameters
  - For example, when using deviation from means we estimate using deviations from average (might be small)
- Random effects:
  - Need to assume that the random effect is uncorrelated with included vairables
- Which model is the right one?
  - For example, test for correlation between residuals and explanatory variables
  - Intuition

Cost Functions

- Similarly one can estimate a cost function:

$$(a)\ \ln c_i = \beta_o + \beta_y \ln y_i + \varepsilon_i \text{ where}$$

c_i are the operating costs of plant (or firm) i

ε_i is an error term (due for example to managerial efficiency)

β_o and β_y are parameters to be estimated

- However, larger firms may be able to obtain raw materials at a lower price:
  - This does not reflect managerial efficiency but contained in ε_i
  - Correlation between error term and included variable
  - OLS estimator may be biased. Need to control for it!

Cost Functions

- If one has data on input prices:

$$(b)\ \ln c_i = \beta_o + \beta_w \ln w_i + \beta_r \ln r_i + \beta_y \ln y_i + \varepsilon_i \text{ where}$$

c_i are the operating costs of plant (or firm) i

w_i is the price of the labour input (wages) of plant (or firm) i

r_i is the price of the capital input of plant (or firm) I

ε_i is an error term (due again for example to managerial efficiencies)

β_o,β_w, β_r and β_y are parameters to be estimated

- From now, however, suppose that companies are identical except wrt size of the output

Regression Analysis

- Assume, for simplicity, that we estimate the model

$$(c)\ \ c_i = \beta_o + \beta_y y_i + \varepsilon_i$$

where β_o is the fixed cost and β_1 is the marginal cost of y

- Then

$$c_i = \hat{c}_i + \hat{\varepsilon}_i$$

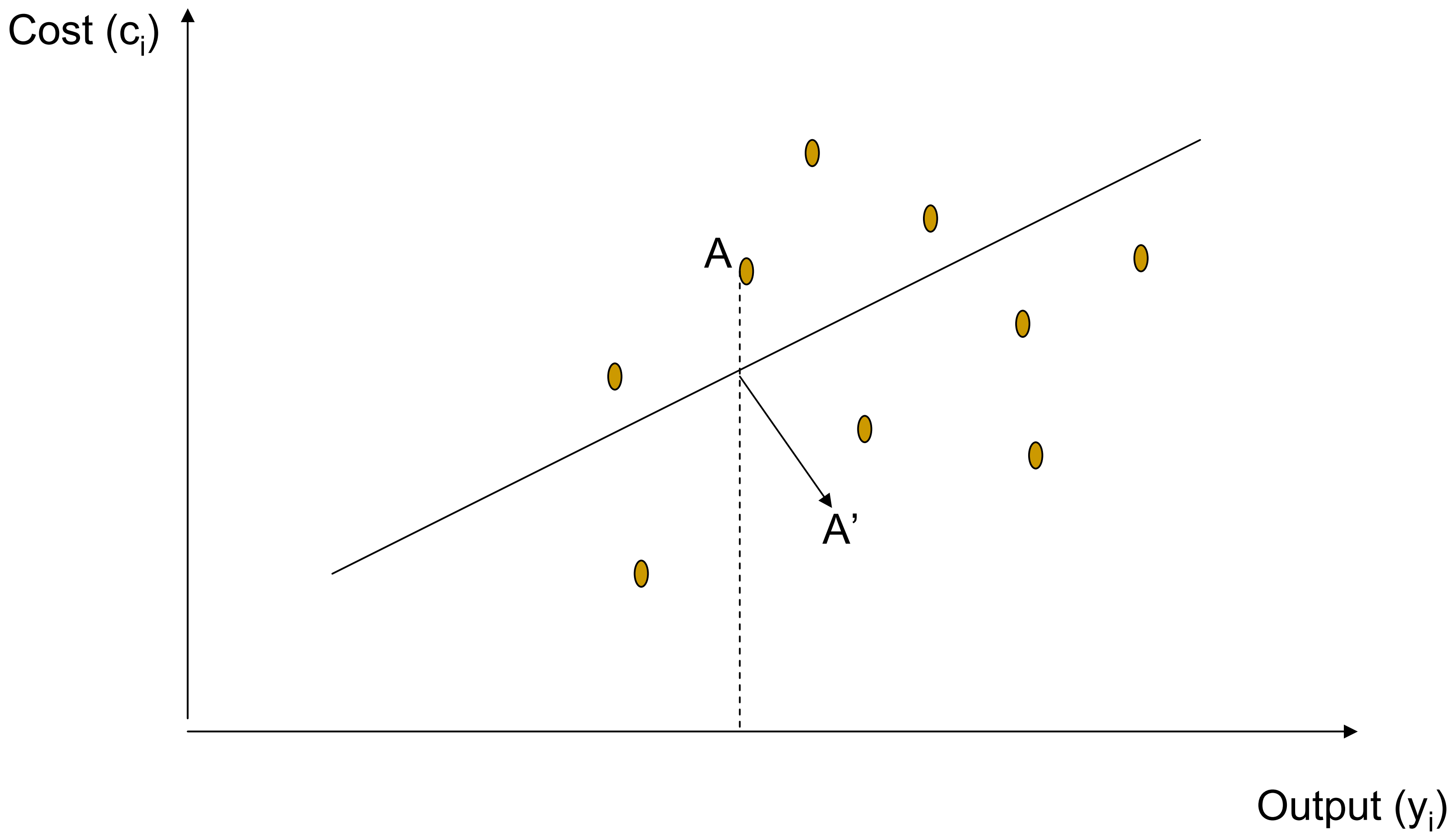
Where $\hat{c}_i$ can be thought as an estimate of the costs of an average efficiency company (*benchmark*) producing the same output

And $\hat{\varepsilon}_i$ estimates the efficiency difference between company i and this average company

The company is more efficiency than average if $\hat{\varepsilon}_i > 0$

The company is less efficiency than average if $\hat{\varepsilon}_i < 0$

Regression Analysis



Extensions

- More than one cost driver:

$$(d)\ \ c_i = \beta_o + \beta_p P_i + \beta_L L_i + \beta_y y_i + \varepsilon_i$$

- More than one year (panel, see Lab):

$$(e)\ \ c_{i,t} = \beta_o + \beta_p P_{i,t} + \beta_L L_{i,t} + \beta_y y_{i,t} + t + \varepsilon_i$$

- Non-linear functional forms? (logs,…)

Index Numbers and TFP

- Instrument to measure change in levels of various related economic variables (RPI,…)
- Here, some price and quantity index numbers relevant to productivity analysis
- Objectives:
  - Measure changes in Total Factor Productivity (TFP)
  - “Aggregate data” into smaller sets of inputs and outputs for DEA and SFA

Price Index Numbers

- In a multi-product setting, how do we measure price and quantity changes, across time or individuals?
- Tornqvist price index P^T, e.g., from s to t is defined as:

$$\ln P_{s,t}^{T} = \sum_{i=1}^{N} \left(\frac{\omega_{i,s} + \omega_{i,t}}{2} \right) \left[\ln p_{i,t} - \ln p_{i,s} \right]$$

where $p_{i,s}$ and $q_{i,s}$ are the price and quantity of i product in s-period

$$\omega_{i,s} = \frac{p_{i,s} q_{i,s}}{\sum_{i=1}^{N} p_{i,s} q_{i,s}}$$ is the value of the i product in period s

- Others include Laspeyres, Paasche, Fisher

Quantity Index Numbers & TFP

- Similar to before we have several possible indices
- Tornqvist quantity index Q^T from s to t is defined as before:

$$\ln Q^T_{s,t} = \sum_{i=1}^{N} \left(\frac{\omega_{i,s} + \omega_{i,t}}{2} \right) \left[\ln q_{i,t} - \ln q_{i,s} \right]$$

- With Tornqvist quantity indices, e.g., compute TFP:

$$\ln TFP_{s,t} = \ln \frac{Output^T_{s,t}}{Input^T_{s,t}} = \ln Output^T_{s,t} - \ln Input^T_{s,t}$$