國科會社會科學研究中心學術研習營

成本隨機邊界模型介紹

王泓仁
中央研究院經濟研究所
hjwang@econ.sinica.edu.tw

國立嘉義大學

June 2-3, 2007

Basics of a Cost Function

- (1) Output, (2) input prices, and (3) production technology are given, producers choose input combinations (both quantity and relative shares) to minimize cost.

- The cost may not be minimized if

  ◇ inputs are not in the appropriate ratio $\Longrightarrow$ allocative inefficiency, and/or

  ◇ input shares are correct but they are not used efficiently in the production $\Longrightarrow$ technical inefficiency.

- technical inefficiency:
  - ◇ output-oriented (OO) tech. ineff. ⇒ easier for production frontiers;
  - ◇ input-oriented (IO) tech. ineff. ⇒ easier for cost frontiers.
- allocative inefficiency:
  - ◇ arising from not using the best input combinations (given prices);
  - ◇ only matters to cost frontiers (not production frontiers);

What We Will Do

- cost frontier models with only IO technical inefficiency (assuming all the firms are allocatively efficient).
  - ◇ estimate as a single equation (very similar to the production frontier model);
  - ◇ estimate using a system of equations (cost function and share equations).

Remark

- Inefficiency should always be traced back to
  - ◇ (1) technical inefficiency, which is defined on the production technology, and
  - ◇ (2) allocative inefficiency, which is defined on the input mix.
- A common mistake is in ignoring the source of cost inefficiency.
  - ◇ only say it is "cost inefficiency"; but why?
  - ◇ will be clear when we see equations

The Cost Frontier Model with IO Technical Inefficiency

- Output, input prices, and production technology are given; choose input combinations to minimize costs.

- $$\min_{x_j} \; \boldsymbol{w}'\boldsymbol{x} \quad \text{s.t.} \quad y = f(\boldsymbol{x}e^{-\eta}),$$

$$\text{FOC:} \qquad \frac{f_j(\boldsymbol{x}e^{-\eta})}{f_1(\boldsymbol{x}e^{-\eta})} = \frac{w_j}{w_1}, \quad j = 2, ..., J,$$

  ◇ $\eta \geq 0$: input-oriented technical inefficiency; the percentage by which the producers could reduce use of all the inputs while producing the given output;

◇ $\boldsymbol{x}$: observed input usage;

◇ $\boldsymbol{x}e^{-\eta} \leq \boldsymbol{x}$: efficient units of input used in the production;

- Define the *cost frontier* (minimum cost, which is unobservable) while ignoring allocative inefficiency: the minimal amount of cost necessary to produce y

$$C^*(\boldsymbol{w}, y) = \sum_j w_j x_j e^{-\eta},$$

- It can be shown that the solution of the model involves the following equations:

$$\ln C^a(\boldsymbol{w}, y, \eta) = \ln C^*(\boldsymbol{w}, y) + \eta + v, \tag{1}$$

$$S_j = \frac{\partial \ln C^*}{\partial w_j} + \zeta_j, \qquad j = 2, \ldots, J. \tag{2}$$

where C^a is the actual, observed cost of the firm, S_j is the cost share of input j, and v is the added statistical error.

- ◇ Note that the IO technical inefficiency (η) is transmitted from the production function to the cost function ($\ln C^a(\cdot)$)!!
- ◇ Use only $J - 1$ share equations because $\sum S_j = 1$ so one of them would be redundant for the estimation purpose.

- Note that (2) contains no new parameters; therefore, the model can be estimated

- ◇ either using the single equation of (1) (easy), or
- ◇ using the system of equations of (1) and (2) (more difficult). System approach gains efficiency.

- The corresponding OLS residuals should skew to the right (positive skewness).

What IF with Allocative Inefficiency

- $\frac{f_j}{f_1} \neq \frac{w_j}{w_1}$: The input *combination* is not optimal in the cost minimization context. Let

  $$\frac{f_j}{f_1} = \frac{w_j}{w_1} \cdot \exp(\xi_j), \qquad \xi_j >=< 0, \quad \text{so} \quad 0 < \exp(\xi_j) \leq 1, \quad \exp(\xi$$

  then input j is over used relative to input 1 if $\exp(\xi_j) < 1$ or $\xi_j < 0$.

Cost Frontier: a Graph

Measures of (In)efficiency

- Since $\ln C^a = \ln C^* + \eta$,

$$\exp(-\eta) = \frac{C^*}{C^a}, \tag{3}$$

therefore $\mathbf{exp(-\boldsymbol{\eta})}$ is the **efficiency score** which measures the minimum cost as a ratio of actual cost. Btw 0 and 1.

- Also, because

$$\eta = \ln C^a - \ln C^* = \ln(C^a/C^*) \tag{4}$$

$$= \ln((C^* + \Delta)/C^*) = \ln(1 + \Delta/C^*), \tag{5}$$

so $\boldsymbol{\eta}$ is the **inefficiency score** for which $100 \times \eta$ is approximately the percentage by which the actual cost exceeds the minimum cost.

The Single-Equation Approach

- Notice that, from the functional point of view, a cost frontier and a production frontier are very similar:

$$\text{(prod frontier:)} \quad \ln y_i = \ln f(\cdot) + v_i - u_i, \quad u_i \geq 0; \tag{6}$$

$$\text{(cost frontier:)} \quad \ln C_i = \ln C^*(\cdot) + v_i + \eta_i, \quad \eta_i \geq 0. \tag{7}$$

- Therefore, **all** the estimation strategies/ methods/ equations introduced earlier can be applied here with some algebraic adjustments.

Parametric Cost Frontier Model: half-normal

$$
\begin{aligned}
\ln C^a &= \ln C^*(\boldsymbol{x}, y) + \eta + v \\
&= \ln C^*(\boldsymbol{x}, y) + \epsilon, \\
\eta &\sim N^+(0,\ \sigma^2), \\
v &\sim N(0,\ \sigma_v^2),
\end{aligned}
$$

$$
L = -\ln(0.5) - \frac{1}{2}\ln(\sigma_v^2 + \sigma^2) + \ln\phi\left(\frac{-\epsilon}{\sqrt{\sigma_v^2 + \sigma^2}}\right) + \ln\Phi\left(\frac{\mu_*}{\sigma_*}\right),
$$

where

$$\mu_* = \frac{\sigma^2 \epsilon}{\sigma_v^2 + \sigma^2}, \quad \sigma_*^2 = \frac{\sigma_v^2 \sigma^2}{\sigma_v^2 + \sigma^2}.$$

- Compared to the half-normal production frontier model, the only difference is in the opposite sign of ϵ.

Estimation Steps

- Choose a functional form for $\ln C^*(\cdot)$, and impose the price homogeneity condition on parameters (and, if applicable, parameter symmetry):
  - ◇ Cobb-Douglas (CD); easy, but inflexible;
  - ◇ Translog (TL); difficult, but more commonly used;
- Choose how to deal with η_i
  - ◇ distribution-free: COLS, thick frontiers, etc.,
  - ◇ distribution-based: half-normal, truncated normal, etc..
- After the model is estimated, check the regularity conditions such as increasing cost in price, etc..

Example of a TL Form (1/2)

$$
\begin{aligned}
\ln C^a =& \ln C^*(\boldsymbol{w}, y) + v + \eta \\
=& \beta_0 + \sum_j \beta_j \ln w_j + \beta_y \ln y + \frac{1}{2} \sum_j \sum_k \beta_{jk} \ln w_j \ln w_k \\
& + \frac{1}{2}\beta_{yy} \ln y \ln y + \sum_j \beta_{jy} \ln w_j \ln y + v + \eta.
\end{aligned} \tag{8}
$$

- Symmetric restrictions and homogenous of degree 1 in input prices requires

$$\beta_{jk} = \beta_{kj}, \qquad \sum_j \beta_j = 1, \qquad \sum_j \beta_{jk} = 0, \qquad \sum_j \beta_{jy} = 0. \tag{9}$$

An Example of a TL Form (2/2)

- Assume $j = 2$, then

$$\ln\left(\frac{C^a}{w_1}\right) = \beta_0 + \beta_y \ln y + \beta_2 \ln\left(\frac{w_2}{w_1}\right) + \frac{1}{2}\beta_{yy}(\ln y)^2 + \frac{1}{2}\beta_{22} \ln\left(\frac{w_2}{w_1}\right)^2 + v + \eta. \quad (10)$$

The System of Equations Approach

- Recall the system of equations:

$$\ln C^a(\boldsymbol{w}, y, \eta) = \ln C^*(\boldsymbol{w}, y) + \eta + v,$$
$$S_j = \frac{\partial \ln C^*}{\partial w_j} + \zeta_j, \qquad j = 2, \ldots, J.$$

- Price homogeneity requires parameter constraints on (1) and (2).

- Since $\ln C^*(\cdot)$ are on both (1) and (2), cross-equation parameter constraints need to be imposed during the estimation

example: 3 variable inputs 1 quasi-fixed input

Estimation Method

- Choose the functional form of $\ln C^*(\cdot)$, and impose price homogeneity (and parameter symmetry if applicable).

- Make distributional assumptions on error terms, such as

$$v \sim N(0, \sigma_v^2), \quad \eta \sim N^+(\mu, \sigma^2), \quad \boldsymbol{\zeta} \sim N(\mathbf{0}, \Sigma). \tag{11}$$

  Also make assumptions of whether $\boldsymbol{\zeta}$ are correlated among themselves, and/or whether $\boldsymbol{\zeta}$ and η are correlated.

- Derive the log-likelihood function, and obtain the estimates via MLE.

- The model statistics (i.e., (in)efficiency index, etc.) are obtained as in the case of the single equation model.

Empirical Example: Efficiency of Commercial Airlines

- Data (all in log):
  - ◇ tc: total cost; y: otuput;
  - ◇ pl, pf, pk: prices of labor, fuel, and capital;
  - ◇ sl, sf, sk: the cost shares of labor, fuel, can capital.
- $\ln C^*(\cdot)$ is assumed to have a translog form;
- pf is used to normalize cost and other prices to maintain price homogeneity.

The Model

$$
\begin{aligned}
tc =& \beta_0 + \beta_1 pl + \beta_2 pk + \beta_y y + [\beta_{11} plpl2 + \beta_{22} pkpk2 + \beta_{yy} yy2] \\
&+ \beta_{12} plpk + \beta_{1y} ply + \beta_{2y} pky + \eta + v, \\
sl =& \beta_1 + \beta_{11} pl + \beta_{12} pk + \beta_{1y} y + \zeta_1, \\
sk =& \beta_2 + \beta_{12} pl + \beta_{22} pk + \beta_{2y} y + \zeta_2.
\end{aligned}
$$

以下內容將以電腦實際操作顯示。

$\eta = 0$: estimate by SURE

- show the estimation and the results;
- use the residuals to perform skewness tests;

SURE results

half-normal, single equation

half-normal, system, no correlation

Show and discuss other cases.