國科會社會科學研究中心學術研習營

生產隨機邊界模型介紹

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

Basic

- 古典生產理論認爲, 平均而言, 生產者皆能將生產技術充分發揮。
- 生產邊界理論則認爲，由於疏忽、經驗不足、管理不當等因素，生產技術常常無法被充分發揮，導致 系統性 (systematic) 的「無效率」的存在。

如何定義與衡量無效率?

- 以生產函數爲著眼點
  - ◇ 要素數量爲外生給定, 產量爲內生;
  - ◇ 可能引起技術無效率 (較簡單);
- 以成本函數爲著眼點
  - ◇ 給定產量及要素價格; 要素使用量爲內生;
  - ◇ 可能引起技術無效率及配置無效率 (較複雜);
- 以利潤函數爲著眼點
  - ◇ 產量與要素使用量皆爲內生;
  - ◇ 可能引起技術無效率 + 配置無效率 (較複雜)。

生產函數

- $y^* = f(x_1, x_2, \ldots, x_n; \boldsymbol{\beta}) \equiv f(\boldsymbol{X}; \boldsymbol{\beta})$

◇ $\boldsymbol{X}$: 生產投入

◇ $f(\cdots)$:「生產技術」

⋆ 例:1麵包 = 0.5個蛋 + 0.5斤麵粉;1蛋糕=2個蛋+1斤麵粉;

◇ $y^* = f(\mathbf{X}; \beta)$ 表示當「充分發揮 生產技術」時的產出與投入的關係。即: 給定投入 $\mathbf{X}$ 與生產技術 $f(\cdot)$時, 所能得到的「最大」產出.

例子:2項投入;3D

$y=f(x_1,x_2;\beta)$描繪的是圓椎體的「表面」。

- 裡面的點?!

例子:2項投入; 縱切面 (給定x_2)

sf model 6

- 古典:(平均而言) 生產活動/行爲皆發生在圖1圓椎的表面, 或圖2的最大生產可能線。
- 生產邊界理論: 生產活動有可能在圖1的圓椎之「內」, 或圖2的 生產可能線之下。亦即, 存在「技術無效率」(technical inefficiency; TI), 使得產量無法達到最大。

兩種角度評估技術無效率

- **用了太多的投入!**生產y_b卻用了x_1^1; if fully efficient, need only $x_1^2 < x_1^1$. $\Longrightarrow$ input-oriented (IO) measure of technical inefficiency.

- **得到太少的產出!**使用x_1竟只得到y_b!If fully efficient, should get $y_B > y_b$. $\Longrightarrow$ output-oriented (OO) measure of technical inefficiency.

IO technical inefficiency

$$y = f(x \cdot e^{-\eta}),\ \eta \geq 0.$$

- Note: $0 < e^{\theta} < 1$ if $\theta < 0$; $e^{\theta} = 1$ if $\theta = 0$; $e^{\theta} > 1$ if $\theta > 0$.

  ◇ Therefore, $0 < e^{-\eta} \leq 1$ given $\eta \geq 0$.

- x:「觀察到」的使用投入 (太浪費);y:「觀察到」的產出;
- $xe^{-\eta}$: 在給定的技術水準下 (i.e., $f(\cdot)$), 使用較少的投入, 就可得到相同的產出。

OO technical inefficiency (較常用)

- $y = y^* \cdot e^{-u}$, where $u \geq 0$ and $y^* = f(x;\beta)$ (古典)
  - ◇ therefore,

$$y = f(x;\beta) \cdot e^{-u}, u \geq 0, \tag{1}$$
$$\Rightarrow y e^{u} = f(x;\beta); \tag{2}$$

若 fully efficient, 使用相同的投入, 可得到比觀察到的更多的產出; $ye^u > y$.

OO technical inefficiency of a production function: more closely

$y = f(x)e^{-u} = y^*e^{-u}$, $u \geq 0$; $e^{-u} = y/y^*$; $-u = \ln(y/y^*)$ so
$u = \ln(y^*/y) = \ln((y + \Delta y)/y) = \ln(1 + \Delta y/y) \approx \Delta y/y$.

Measures of (In)Efficiency

- **efficiency score**:

$$y = f(x)e^{-u} = y^* e^{-u}, u \geq 0; \Longrightarrow e^{-u} = y/y^*;$$

so e^{-u} measures the percent of potential output actually achieved; a measure of **technical efficiency**;

- **inefficiency score**;

$$-u = \ln(y/y^*),$$

$$u = \ln(y^*/y) = \ln((y + \Delta y)/y) = \ln(1 + \Delta y/y) \approx \Delta y/y;$$

so u measures the percent of output loss (when multiplied by 100) due to technical inefficiency.

估計模型之前...

- A crucial question: Do the samples share the same production technology?
  - ◇ The inefficiency is measured as the distance between the actual output and the optimal output.
  - ◇ Optimal output is measured conditional on the production technology; that is, the production technology is the same for all individuals.
  - ◇ If they do not share the same technology, the estimates of inefficiency is misleading.
- Bad examples: pooling China and US banks together; pooling firms from very different industries.

模型估計的目的

- 探討生產技術 (i.e., the β coefficients);
- 瞭解整體而言, 無效率是否嚴重 (i.e., $E(u_i)$ or $E(e^{-u_i})$);
- 分析導致無效率的因素 (i.e., what determines u_i);
- 比較生產者間的生產效率, 是否顯著不同 (i.e., $E(u_i)$ vs. $E(u_j)$);
- 將生產者的效率予以排序。

估計方法(OO;production frontier)

- Two goals:
  - ◇ obtain consistent coefficients (β),
  - ◇ obtain *observation-specific* (in)efficiency index.

$$y_i = f(x_i;\beta)e^{-u_i}, \Rightarrow \ln y_i = \ln f(x_i;\beta) - u_i; \qquad (3)$$

- 如何處理 u_i?
  - ◇ distribution free approach:
    - ⋆ COLS (corrected OLS);

⋆ thick frontier approach; (will skip)

◇ parametric approach (based on distribution assumptions).

⋆ half normal, truncated normal, etc.;

Distribution-free approach: COLS

- Does not require (strict) distribution assumptions on u_i.
- Estimate the model by OLS, and then shift the intercept to have the interpretation of a stochastic frontier.

◇ **step 1:** Want to estimate

$$y_i = \beta_0 + \beta_1 x_i - u_i, \quad u_i \geq 0, \tag{4}$$

but estimate instead the following by OLS

$$y_i = \beta_0^* + \beta_1^* x_i + v_i, \quad v_i \sim N(0, \sigma^2). \tag{5}$$

$\hat{\beta}_1^*$: consistent estimate of β_1; $\hat{\beta}_0^*$: *not* a consistent estimate of β_0.

◇ **step 2:** shift the regression line *upward* by the amount of $\max\{\hat{v}_i\}$ (to bound all the data below) by adjusting the intercept:

$$\hat{v}_i - \max\{\hat{v}_i\} = \ln y_i - \underbrace{\left\{\left[\hat{\beta}_0^* + \max\{\hat{v}_i\}\right] + \tilde{\boldsymbol{x}}_i'\hat{\tilde{\boldsymbol{\beta}}}^*\right\}}_{\text{estimated frontier function}} \leq 0, \tag{6}$$

and that

$$\hat{u}_i \equiv -(\hat{v}_i - \max\{\hat{v}_i\}) \geq 0, \tag{7}$$

$\hat{u}_i$: the estimated inefficiency index; technical efficiency of each observation: $\widehat{TE}_i = \exp(-\hat{u}_i)$.

Comments on COLS

- Advantages:
  - ◇ easy to estimate; no special software needed;
- Disadvantages:
  - ◇ COLS is a shifted mean regression $\Longrightarrow$ difference between the efficient frontier producer and the mean producer only affects the intercept but not slope coefficients;
  - ◇ the frontier function is deterministic $\Longrightarrow$ (1) the randomness of the model comes entirely from inefficiency variation; (2) deviations from the estimated frontier are entirely attributed to inefficiency; (3) estimates of inefficiency would be sensitive to outliers.

Parametric Approach: Setup

$$\ln y_i = \ln f(x_i;\beta) - u_i + v_i, \quad \text{where } v_i \sim N(0,\sigma^2), \tag{8}$$

- impose distributional assumption on u_i, and derive the likelihood function;
- obtain estimates of the coefficients by MLE;
- obtain *observation-specific* measure of technical inefficiency (tricky!).

Parametric Approach: u is Half-Normal

$$\ln y_i = \ln y_i^* - u_i, \tag{9}$$

$$\ln y_i^* = \boldsymbol{x}_i \boldsymbol{\beta} + v_i, \quad \text{(the efficient stochastic frontier)}, \tag{10}$$

$$u_i \sim N^+(0,\, \sigma^2), \qquad v_i \sim N(0,\, \sigma_v^2), \tag{11}$$

Or

$$\ln y_i = \boldsymbol{x}_i \boldsymbol{\beta} + \epsilon_i; \quad \epsilon_i = v_i - u_i, \tag{12}$$

(half normal:) Density Plot

(half normal:) Log-Likelihood Function

$$L_i = -\ln(0.5) - \frac{1}{2}\ln(\sigma_v^2 + \sigma^2) + \ln\phi\left(\frac{\epsilon_i}{\sqrt{\sigma_v^2+\sigma^2}}\right) + \ln\Phi\left(\frac{\mu_{*i}}{\sigma_*}\right), \tag{13}$$

where

$$\mu_{*i} = \frac{-\sigma^2\epsilon_i}{\sigma_v^2+\sigma^2}, \tag{14}$$

$$\sigma_*^2 = \frac{\sigma_v^2\sigma^2}{\sigma_v^2+\sigma^2}. \tag{15}$$

(half normal:) Observation-Specific Measure of Inefficiency Index

- Recall that, for a production frontier model, inefficiency is measured by $E(u_i)$ $(= \Delta y/y)$: percent of output loss.

- How? $u \sim N^+(0, \sigma^2)$, we obtain $\hat{\sigma}^2$, but then $E(u)$ is a single value, not observation-specific?!

- Jondrow et al. (1982) solve the problem by using conditional expectation:

◇ $E(u_i|\hat{\epsilon}_i) = E(u_i|y_i - X_i\hat{\beta})$,

◇ since $y_i - X_i\hat{\beta}$ is observation-specific, so would be $E(u_i|\hat{\epsilon}_i)$.

$$E(u_i|\epsilon_i) = \frac{\sigma_* \phi(\frac{\mu_{*i}}{\sigma_*})}{\Phi\left(\frac{\mu_{*i}}{\sigma_*}\right)} + \mu_{*i}. \tag{16}$$

- the lower bound (L_i) and the upper bound (U_i) of a $(1-\alpha)100\%$ confidence interval:

$$L_i = \mu_{*i} + \Phi^{-1}\left\{1 - \left(1 - \frac{\alpha}{2}\right)\left[1 - \Phi\left(-\frac{\mu_{*i}}{\sigma_*}\right)\right]\right\}\sigma_*, \tag{17}$$

$$U_i = \mu_{*i} + \Phi^{-1}\left\{1 - \frac{\alpha}{2}\left[1 - \Phi\left(-\frac{\mu_{*i}}{\sigma_*}\right)\right]\right\}\sigma_*, \tag{18}$$

(half normal:) Observation-Specific Measure of Efficiency Index (popular)

- Recall that, for a production frontier model, efficiency is measured by $E(\exp(-u_i))$ $(= y/y^*)$: percent of potential output achieved. Close to 1, more efficient.

- Battese and Coelli (1988), following Jondrow et al. (1982), suggests a conditional expectation measure:

$$E\left[\exp(-u_i)|\epsilon_i\right] = \exp\left(-\mu_{*i} + \frac{1}{2}\sigma_*^2\right) \frac{\Phi\left(\frac{\mu_{*i}}{\sigma_*} - \sigma_*\right)}{\Phi\left(\frac{\mu_{*i}}{\sigma_*}\right)}. \quad (19)$$

- Confidence Intervals:

$$\mathcal{L}_i = \exp(-U_i), \tag{20}$$

$$\mathcal{U}_i = \exp(-L_i). \tag{21}$$

(half-normal:) exogenous determinants

- Approach 1: After obtaining $E(u_i|\epsilon_i)$, do a Logit regression of $E(u_i|\epsilon_i)$ on exogenous determinants $\mathbf{z}_i$. But this approach has serious problems:
  - ◇ the iid assumption of u_i is violated;
  - ◇ Wang and Schmidt (2002) show that the bias is significant.
- Approach 2: (Ford et al. 1993) $u \sim N^+(0, \sigma^2 = g(\mathbf{Z}_i\delta)) \equiv N^+(0, \exp(\mathbf{Z}_i\delta))$
  - ◇ be careful of the interpretation: δ is not marginal effect, although the sign is the same as the marginal effect's sign.

(half-normal:) test the one-sided error (u_i)

- Since $u_i \sim N^+(0, \sigma^2)$, it amounts to test $H_0 : \sigma^2 = 0$. An LR test statistic of $-2[L(H_0) - L(H_1)]$ was suggested, where $L(H_0)$ and $L(H_1)$ are log-likelihood values of the restricted model (OLS) and the unrestricted model (SF).

- Problem: The null ($\sigma^2 = 0$) is on the boundary of the parameter space, and usual distributions do not apply.

- Coelli (1995) shows that the LR the test has a mix chi-square distribution, with critical values given in Kodde and Palm (1986).

(half normal:) Comments

- Advantages:
  - ◇ easier to estimate (has only one parameter in u's distribution);
- Disadvantages:
  - ◇ inflexible because has only one parameter to characterize the distribution;
  - ◇ the mode is at 0 (the full efficiency level).

(half normal:) Remark 1

- Before you begin, it is always a good idea to check the skewness of the model's corresponding OLS residuals.
- (Schmidt and Lin, 1984) for a stochastic production frontier model, the OLS residuals should skew to the left (i.e., **negative** skewness) (recall $v_i - u_i$).
- $$\sqrt{b_1} = \frac{m_3}{m_2\sqrt{m_2}}, \tag{22}$$

  m_2 and m_3: the second and the third sample moments of the OLS residuals.

(half-normal:) Remark 2

- An often-reported but sometimes mis-used statistic:

$$\gamma = \frac{\sigma^2}{\sigma^2 + \sigma_v^2}. \tag{23}$$

- For half-normal model, the gamma may be used to test for the significance of the u_i term. (But could be used for truncated normal model; later.)

- It is often interpreted as the share of model variance coming from inefficiency. This is WRONG, because $var(u) = (1 - 2/\pi)\sigma^2 \neq \sigma^2$.

Parametric Approach: u is Truncated Normal

$$\ln y_i = \boldsymbol{x}_i\boldsymbol{\beta} + \epsilon_i, \tag{24}$$

$$\epsilon_i = v_i - u_i, \tag{25}$$

$$u_i \sim N^+(\mu,\, \sigma^2), \tag{26}$$

$$v_i \sim N(0,\, \sigma_v^2). \tag{27}$$

- The functional form is more flexible than the half-normal distribution.
- The mode of the distribution is not necessarily 0.

(truncated normal:) Density Plot

Figure 1: Density Plot of Truncated-Normal Distributions

(truncated normal:) log-likelihood function

$$L_i = -\frac{1}{2}\ln(\sigma_v^2+\sigma^2)+\ln\phi\left(\frac{\mu+\epsilon_i}{\sqrt{\sigma_v^2+\sigma^2}}\right)+\ln\Phi\left(\frac{\mu_{*i}}{\sigma_*}\right)-\ln\Phi\left(\frac{\mu}{\sigma}\right),$$

where

$$\mu_{*i} = \frac{\sigma_v^2\mu - \sigma^2\epsilon_i}{\sigma_v^2+\sigma^2},$$

$$\sigma_*^2 = \frac{\sigma_v^2\sigma^2}{\sigma_v^2+\sigma^2}.$$

(truncated normal:) Statistics

the observation-specific measure of inefficiency and the confidence intervals

- *skip; similar to those of the half-normal distribution; see book*

(truncated normal:) Exogenous Determinants

- Given $u_1 \sim N^+(\mu, \sigma^2)$, how to account for the exogenous determinants of inefficiency?
- Battese and Coelli (1995): $u_i \sim N^+(\mu = \mathbf{Z}_i\delta, \sigma^2)$;
- Problem: Why $\mathbf{Z}_i$ enters through μ but not σ^2 (Wang, 2002)?

(truncated normal:) Problem 1: Insufficient Parameterization

The moments of $\boldsymbol{u}$ are functions of both $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}^2$.

$$
\begin{aligned}
E(u) &= f(\mu,\ \sigma) = \mu + \sigma \left[\frac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} \right], \\
V(u) &= g(\mu,\ \sigma) = \sigma^2 \left[1 - \frac{\mu}{\sigma} \left[\frac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} \right] - \left[\frac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} \right]^2 \right].
\end{aligned}
$$

- There is no theoretical justification why exogenous factors $\boldsymbol{z_{it}}$ affect the moments of $\boldsymbol{u}$ through $\boldsymbol{\mu}$ but not $\boldsymbol{\sigma}$.

- Different elements in $\boldsymbol{z_{it}}$ may influence $\boldsymbol{u}$ through different channels.

(truncated normal:) Problem 2: Efficiency Effects Are Necessarily Monotonic

- If only μ is parameterized, the relationship between $E(u)$ and elements of z_{it} are forced to be monotonic.

- Many economic relationships, however, are likely to be non-monotonic. For instances, the income tax rate and government revenues; the wage rate and labor supply; ages and labor efficiency; financial ratios (debt-asset ratio, interest-income ratio) and firm performances.

- **To sum up:** When z_{it} is only allowed in μ, the exogenous factors' effects on u may not be fully accounted for, and their abilities to explain the model are compromised and restricted.

(truncated normal:) The Improvement (Wang 2002)

- Make the model more flexible by allowing $\boldsymbol{z_{it}}$ to affect $\boldsymbol{u}$ through both $\boldsymbol{\mu}$ and $\boldsymbol{\sigma^2}$.

$$\mu = \mathbf{Z}_i\delta, \quad \sigma^2 = \exp(\mathbf{Z}_i\gamma) \tag{28}$$

- This added flexibility has the benefit of accommodating non-monotonic efficiency effects of $\boldsymbol{z_{it}}$ on $\boldsymbol{E(u)}$. The marginal effect of the kth element of $\boldsymbol{z_{it}}$ is:

$$\begin{aligned}
\frac{\partial \mathrm{E}(u_{i\,t})}{\partial z[k]} &= \delta[k]\left[1-\frac{\mu}{\sigma}\left[\frac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)}\right]-\left[\frac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)}\right]^2\right] \\
&\quad + \frac{\gamma[k]}{2}\left[\left(\sigma+\mu\frac{\mu}{\sigma}\right)\left[\frac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)}\right]+\mu\left[\frac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)}\right]^2\right].
\end{aligned}$$

(truncated normal:) An Empirical Example

- **Data** Outputs and inputs of Indian rice farmers. Unbalanced panel of 34 farmers, with a total number of observations equal to 271.

- **Variables** Same as Battese and Coelli (1995) and Coelli and Battese (1996).

y_{it}: $\ln(Y_{it})$;

$\mathbf{x}_{it}$: $\{\ln(Land_{it}),\ \ln(PILand_{it}),\ \ln(Labor_{it}),\ \ln(Bullock_{it}),\ \ln[\text{Max}(Cost_{it}, 1 - D_{it})],\ Year_{it}\}$;

$\mathbf{z}_{it}$: $\{Age_{it},\ Schooling_{it},\ Year_{it}\}$,

(truncated normal:) Marginal Effects of Age on Inefficiency $E(u_{it})$

Figure 2: (i): Battese and Coelli (1995); (ii): Wang (2002)

(truncated normal:) Remark 1

- Recall the γ statistic:

$$\gamma = \frac{\sigma^2}{\sigma^2 + \sigma_v^2}. \tag{29}$$

- Is σ^2 related to the variance of u_i?

◇ $u_i \sim N^+(\mu, \sigma^2)$, so $var(u_i)$ is a function of both μ and σ^2.

- It conveys no useful information for a truncated normal model! So, do not use it!!

Other distributional assumptions

- exponential distribution:
  - ◇ Has only 1 parameter, easy to estimate, very similar to the half-normal distribution.
- Gamma distribution (Greene):
  - ◇ Difficult to estimate (flat likelihood surface).

Scaling Property Model (Wang and Schmidt 2002)

- An variant of the truncated normal model.
- $u_i \sim h(\mathbf{Z_i}, \delta) \times u^*$:
  - $\diamond$ $u^* \geq 0$: *basic distribution*, common to all observation and is not depend on $\mathbf{Z_i}$. It can take any functional form (such as truncated normal)
  - $\diamond$ $h(\mathbf{Z_i}, \delta) \geq 0$: *scaling function*. For example, $h(\cdot) = \exp(\mathbf{Z_i}\delta)$
- The model implies that the random variable u_i follows a common distribution u^*, but each is weighted by a different,

observation-specific scale of $h(\cdot)$.

- The *shape* of the distribution of u_i is the same for all firms. The scaling function $h(\cdot)$ essentially stretches or shrinks the horizontal axis, so that the scale of the distribution of u_i changes but its underlying shape does not.

- Another advantage of this distribution:

$$\frac{\partial \ln u_i}{\partial z[k]} = \delta[k] \tag{30}$$

So, the coefficient has an intuitive interpretation (the effect on the change rate of inefficiency).

Examples

- data: dairy farm production.
  - ◇ `ly`: output: the log of milk production,
  - ◇ `llabor`: labor hours,
  - ◇ `lfeed`: feed,
  - ◇ `lcattle`: number of cows,
  - ◇ `lland`: the farm's land size
  - ◇ `comp`: computer-related expenses to total expenses
- Software: by Hung-Jen Wang in Stata

以下内容將以電腦實際操作顯示。

(examples:) OLS

(examples:) OLS residuals: graph

(examples:) OLS residuals

(examples:) COLS

(examples:) COLS Efficiency Score

(examples:) Half-Normal Model

(examples:) Half-Normal: graph

(examples:) truncated-normal

(examples:) truncated-normal: graph

(examples:) truncated-normal: marginal effects

Do different distributions matter?

Do different distributions matter? Summary Statistics

Do different distributions matter? Correlation Coefficients

Do different distributions matter? Kendall's τ

- τ: $100(1-\tau)/2$ gives the probability that a randomly chosen pair of observations have opposite rankings in the two series.