第 33 卷第 6 期
2016 年 6 月

统计研究
Statistical Research

Vol. 33, No. 6
Jun. 2016

面板数据随机前沿分析的研究综述

边文龙　王向楠

内容提要: 近年来,面板数据随机前沿分析(SFA)越来越多地被用于测算各类决策单位的效率,取得了很多成果,但是国内外实证研究文献也存在过度依赖几种假设严格的模型和不注重模型局限性的问题。本文在统一的计量框架下,对面板 SFA 模型的发展研究进行了系统的梳理总结。本文将相关模型分为效率不随时间变化的模型和效率可随时间变化两大类,每一类又根据是否对效率项的分布做出假设分为有分布假设的模型和无分布假设的模型。在明确和比较不同模型的假设、估计过程和局限性的基础上,对面板 SFA 模型的应用提出了建议。

关键词: 随机前沿分析; 面板数据; 效率; 研究综述
DOI: 10.19343/j.cnki.11-1302/c.2016.06.002
中图分类号: C81　　文献标识码: A　　文章编号: 1002-4565(2016)06-0013-08

A Literature Review on the Stochastic Frontier Analysis in Panel Data

Bian Wenlong & Wang Xiangnan

Abstract: In recent years, Stochastic Frontier Analysis (SFA) in panel data is being increasingly used to measure the efficiency of all forms of decision making units' efficiency and obtaining many research achievements. However, most of these papers rely heavily on several models with strict assumptions and pay insufficient attention to the availability of these models. This paper summarizes the development of Stochastic Frontier Analysis in panel data under a unified econometric framework systematically. We classify these models into two parts based on whether efficiency term changing with time or not. In each part, we further divide the models into two parts according to whether the efficiency term and random error term having distributional assumptions. We emphasize and compare the assumptions, estimation procedures and availability of each model in order to provide potential and helpful suggestions for improving the methods employed in empirical research.

Key words: Stochastic Frontier Analysis; Panel Data; Efficiency; Literature Review

一、引言

效率反映了一个决策单位的综合竞争力,效率研究能够为单位的管理者和股东、行业监管者和宏观政策制定者的决策提供有价值的信息。测算决策单位效率的思想可以追溯到 Koopmas(1951)[1] 和 Debreu(1951)[2] 的研究,Farrell(1957)[3] 引入了前沿分析来测量效率,该方法与微观经济理论中的生产函数、成本函数、利润函数等最优化目标相一致,所以被认为是研究决策单位效率的客观合理的方法。Meeusen 和 van den Broeck(1977)①与 Aigned 等(1977)②最先在截面数据中建立了随机前沿分析(Stochastic Frontier Analysis,SFA)模型,此后,SFA 的计量模型得到了迅速发展,同时也不断发展成为测算各类决策单位效率的有效工具。

与另一种效率研究的常用工具——基于线性规划的数据包络分析方法(data envelop analysis,DEA)相比,SFA 的主要优点在于考虑了由测量误差等因素造成的随机误差,避免了将这些随机误差成份不恰当地计入到效率项之中③。而 SFA 的主要缺点是,研究者一般需要对效率项的分布做出先验的假

① Meeusen W, van den Broeck J. Efficiency estimation form Cobb-Douglas production function with composed errors [J]. International Economic Review, 1977, 18(2): 435-444.

② Aigner D, Lovell C, Schmidt P. Formation and estimation of stochastic frontier production function models [J]. Journal of Econometrics, 1977, 6(1): 21-37.

③ 由于测量误差和不确定性经济环境的问题更可能出现在发展中国家和转型经济国家,所以 SFA 等参数方法较之 DEA 等非参数方法的优势在研究发展中国家和转型经济国家的效率问题时更大。

设，如果假设不当，将造成效率值估计的偏误。然而，基于面板数据的一些 SFA 模型可以克服这个不足，即面板数据 SFA 可以不对效率项的分布做出假设，而且能够允许效率项与模型中的投入产出项之间存在相关性。自 Pitt 和 Lee(1981)[4] 首次将 SFA 应用到面板数据以来，面板数据 SFA 逐渐代替了截面数据 SFA，成为了 SFA 理论和应用研究的主流。

在计量模型迅速发展的同时，运用 SFA 衡量效率的实证文献也不断出现。但是，我们发现，无论是国外还是国内运用 SFA 的文章，都存在着过度依赖几种假设严格的模型和不注重模型的局限性等问题。本文在统一的计量框架下，对面板数据 SFA 模型进行系统的梳理，着重分析每一种模型的假设、估计过程和局限性。同时，本文希望为实证研究者在选择 SFA 的计量方法上提供参考，从而在一定程度上规范未来 SFA 领域的实证研究。

本文结构安排如下：第二部分介绍面板数据 SFA 模型的分类。第三和第四部分分别梳理分析效率不随时间变化的 SFA 模型和效率可随时间变化的 SFA 模型。第五部分梳理分析研究效率影响因素的 SFA 模型。第六部分总结全文并对今后实证研究者选择 SFA 模型提出建议。

二、面板数据 SFA 模型的分类

假设决策单位 i 在时期 t 的生产函数为：

$$Y_{it} = f(X_{it},\beta)\, e^{\varepsilon_{it}} \tag{1}$$

其中，$\varepsilon_{it} = -u_{it} + v_{it}$。$Y_{it}$ 表示实际产出，X_{it} 表示影响产出的 K 个投入，u_{it} 表示无效率项，$u_{it} \geqslant 0$，v_{it} 表示不受决策单位控制的随机因素。$f(X_{it},\beta)\, e^{v_{it}}$ 代表生产前沿，它表示在给定投入下产出的最大值。对式(1)两边同时取对数，可得：

$$\ln(Y_{it}) = \ln f(X_{it},\beta) + \varepsilon_{it} = \ln f(X_{it},\beta) - u_{it} + v_{it} \tag{2}$$

计量模型一般假设变量之间是“线性”关系，即存在 $\ln f(X_{it},\beta) = (\ln X_{it})'\beta$，从而得到：

$$y_{it} = x'_{it}\beta + \varepsilon_{it} = x'_{it}\beta - u_{it} + v_{it} \tag{3}$$

其中，y_{it} 表示产出 Y_{it} 的对数，$x_{it}(K\times 1)$ 表示投入 X_{it} 的对数。

根据 Battese 和 Coelli(1988)① 对“效率”的定义：

$$TE_{it} = \frac{E(Y_{it} \mid U_{it}, X_{it}, t = 1,2,\cdots,T)}{E(Y_{it} \mid U_{it} = 0, X_{it}, t = 1,2,\cdots,T)}$$

$$= \frac{\exp(x'_{it}\beta - u_{it})}{\exp(x'_{it}\beta)} = \exp(-u_{it}) \tag{4}$$

即效率被定义为在相同的投入下，实际产出与完全有效产出的比值。比如，$TE_{it} = 0.90$ 表示，与完全有效的决策单位相比，决策单位 i 在时期 t 可以产生 90% 的产出。前沿分析方法就是按照一定的标准构造一个效率前沿面，再使用决策单位与该前沿面的接近程度来测量效率。从效率的定义中可以看出，它测量的是一定投入、产出下决策单位之间的“相对效率”而不是“绝对效率”。

总的来说，面板数据 SFA 模型可以分为两类：效率不随时间变化的模型和效率可随时间变化的模型。每一类又可以根据是否对效率做出分布假设进一步分为：有分布假设的模型和无分布假设的模型。除了得到效率的估计值和排名之外，研究者经常对效率的“影响因素”感兴趣，我们将其定义为第 5 类模型。表 1 列示了面板数据 SFA 的主要模型。

表 1　面板数据 SFA 模型的分类

分类标准		模型
估计效率值	效率不随时间变化、无分布假设	Schmidt 和 Sickles(1984)
	效率不随时间变化、有分布假设	Pitt 和 Lee(1981)、Battese 和 Coelli(1988)
	效率可随时间变化、无分布假设	Cornwell 等(1990)、Lee 和 Schmidt(1993)
	效率可随时间变化、有分布假设	Kumbhakar(1990)、Battese 和 Coelli(1992)、Greene(2004，2005)、Wang 和 Ho(2010)
研究效率的影响因素		Battese 和 Coelli(1995)、Wang(2002)

三、效率不随时间变化的 SFA 模型

对于效率不随时间变化的模型，即假设 $u_{it} = u_i$，那么，式(3)变为：

$$y_{it} = x'_{it}\beta + \varepsilon_{it} = x'_{it}\beta - u_i + v_{it} \tag{5}$$

Pitt 和 Lee(1981)首次提出了面板数据 SFA 模型，并假设：(a) u_i 在不同的 i 之间服从相互独立的半正态分布，即 $u_i \sim N^+(0,\sigma_u^2)$，则 u_i 的概率密度函数为 $f(u_i) = \frac{2}{\sqrt{2\pi}\sigma_u}\exp(-u^2/2\sigma_u^2)$；(b) v_{it} 在 i 和 t 之间服从相互独立的正态分布 $N(0,\sigma_v^2)$，则 u_i 的概率密度函数为 $g(v_{it}) = \frac{1}{\sqrt{2\pi}\sigma_v}\exp(-v_{it}^2/2\sigma_v^2)$；(c)

① Battese G，Coelli T. Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data [J]. Journal of Econometrics，1988，38(3)：387－399.

u_i 和 v_{it} 之间相互独立; (4) u_i、v_{it} 与 x_{it} 之间不相关。根据以上这些假设,可以得到 ε_i 的概率密度函数,如下:

$$f(\varepsilon_i) = \int_0^\infty \prod_{t=1}^{T} g(\varepsilon_{it} + u_i) f(u_i)\,\mathrm{d}u \tag{6}$$

其中, $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{iT})$, 相应的对数似然函数为:

$$\log L = N\ln 2 - \frac{NT}{2}\ln(2\pi) - \frac{N(T-1)}{2}\ln\sigma_v^2 - \frac{N}{2}\ln(\sigma_v^2 + T\sigma_u^2) - \frac{1}{2\sigma_v^2}\sum_{i=1}^{N}\varepsilon'_i A\varepsilon_i + \sum_{i=1}^{N}\log\left(1 - \Phi\left(\frac{\sigma_u}{\sigma_v(\sigma_v^2 + T\sigma_u^2)^{0.5}}\sum_{t=1}^{T}\varepsilon_{it}\right)\right) \tag{7}$$

其中, $l_n = (1, 1, \cdots, 1)'_{N\times 1}$, $A = I_T - (\sigma_u^2/\sigma_v^2 + T\sigma_u^2)\, l_T l'_T$, $l_T = (1, 1, \cdots, 1)'_{T\times 1}$。通过极大似然估计可以得到参数 β、σ_u^2 和 σ_v^2 的一致估计,并得到残差 $\hat{\varepsilon}_{it}$。根据式(4) 中效率的定义,决策单位 i 对应的效率为 $\exp(-u_i)$。由于无法从 $\hat{\varepsilon}_{it}$ 中分离出 u_i 的估计值 $\hat{u}_i$, 因此 Pitt 和 Lee(1981) 无法得到每个决策单位的效率值,他们只得到了行业内所有决策单位效率值的均值 $E(\exp(-u_i)) = 2\exp(\sigma_u^2/2)(1 - \Phi(\sigma_u))$ 的估计 $2\exp(\hat{\sigma}_u^2/2)(1 - \Phi(\hat{\sigma}_u))$, 其中, $\hat{\sigma}_u^2$ 是 σ_u^2 的极大似然估计。

Jondrow 等(1982)[5] 基于截面数据的 SFA 模型:

$$y_i = x'_i\beta + \varepsilon_i = x'_i\beta - u_i + v_i \tag{8}$$

首次提出从 ε_i 中分离出 u_i 的方法。Jondrow 等(1982) 假设 u_i 在不同的 i 之间服从相互独立的半正态分布 $N^+(0, \sigma_u^2)$, v_i 服从相互独立的正态分布 $N(, \sigma_v^2)$, u_i 和 v_i 之间相互独立。根据 u_i 和 v_i 的联合分布 $f(u_i, v_i)$, 可以得到 u_i 和 ε_i 的边缘分布 $f(u_i, \varepsilon_i)$, 进而得到 u_i 相对于 ε_i 的条件分布 $f(u_i \mid \varepsilon_i)$。Jondrow 等(1982) 证明了 $f(u_i \mid \varepsilon_i)$ 恰好服从从 0 处截断的正态分布 $N^+(\mu_{i*}, \sigma_*^2)$, 其中, $\mu_{i*} = -\frac{\sigma_u^2\varepsilon_i}{\sigma^2}$, $\sigma_*^2 = \frac{\sigma_u^2\sigma_v^2}{\sigma^2}$, $\sigma^2 = \sigma_u^2 + \sigma_v^2$。Jondrow 等(1982) 进而提出用条件期望 $E(u_i \mid \varepsilon_i) = \mu_{i*} + \sigma_*^2 \frac{f(-\mu_{i*}/\sigma_*)}{1 - \Phi(-\mu_{i*}/\sigma_*)}$, 或者用条件分布的众数 $M(u_i \mid \varepsilon_i) = \begin{cases} -\varepsilon_i \frac{\sigma_u^2}{\sigma^2} & (\varepsilon_i \leqslant 0) \\ 0 & (\varepsilon_i > 0) \end{cases}$ 来表示效率。通过极大似然估计得到 u_{i*} 和 σ_* 的估计值,进而得到效率的估计值。

Battese 和 Coelli(1988) 扩展了 Pitt 和 Lee(1981) 的模型,他们假设 u_i 服从从 0 处截断的正态分布 $N^+(\mu, \sigma_u^2)$, 相应的概率密度函数为:

$$f(u_i) = \frac{\exp(-0.5(u_i - \mu)^2/\sigma_u^2)}{2\pi^{0.5}\sigma_u(1 - \Phi(-\mu/\sigma_u))} \tag{9}$$

Battese 和 Coelli(1988) 模型的其他假设与 Pitt 和 Lee(1981) 模型相同。当 $\mu = 0$ 时,从 0 处截断的正态分布退化为半正态分布,所以 Pitt 和 Lee(1981) 模型可以看做 Battese 和 Coelli(1988) 模型在 $\mu = 0$ 的一个特例。同样运用极大似然估计,可以得到参数 β、μ、σ_u^2 和 σ_v^2 的一致估计。与 Pitt 和 Lee(1981) 不同的是, Battese 和 Coelli(1988) 采用 Jondrow 等(1982) 的估计方法,利用:

$$E(\exp(-u_i \mid \varepsilon_i)) = \frac{1 - \Phi(\sigma_*^2 - \mu_{i*}/\sigma_*)}{1 - \Phi(-\mu_{i*}/\sigma_*)}\exp\left(-\mu_{i*} + \frac{1}{2}\sigma_*^2\right)$$

表示决策单位 i 的效率。其中:

$$\mu_{i*} = \left(-\sigma_u^2 T^{-1}\sum_{t=1}^{T}\varepsilon_{it} + T^{-1}\mu\sigma_v^2\right)(\sigma_u^2 + T^{-1}\sigma_v^2)^{-1} \tag{10}$$

$$\sigma_*^2 = \sigma_u^2\sigma_v^2(\sigma_v^2 + T\sigma_u^2)^{-1} \tag{11}$$

通过代入 ε_{it}、μ、σ_u^2 和 σ_v^2 的极大似然估计值,可以得到效率的估计值。

Schmidt 和 Sickles(1984)[6] 指出,假设 u_i 和 v_{it} 服从特定的分布, u_i 和 v_{it} 之间相互独立, u_i、v_{it} 与 x_{it} 之间不相关,都是很严格的。在效率不随时间变化的假设下,决策单位 i 很可能知道自身的效率,从而影响当期投入的选择,即 x_{it} 和 u_i 之间很可能是相关的。一旦 x_{it} 和 u_i 之间存在相关性,效率的估计将存在偏误。当 $u_{it} = u_i$ 时,式(3) 变为①:

$$y_{it} = \alpha + x'_{it}\beta - u_i + v_{it} = x'_{it}\beta + \alpha_i + v_{it} \tag{12}$$

式(12) 本质上就是传统的面板数据模型,运用组内估计量可以得到 β 的一致估计 $\hat{\beta}$, 进而得到 $\hat{\alpha}_i = \bar{y}_i - \bar{x}'_i\hat{\beta}$。令 $\hat{\alpha} = max_i(\hat{\alpha}_i)$, 则 $\hat{u}_i = \hat{\alpha} - \hat{\alpha}_i$, 效率的估计为 $\exp(-\hat{u}_i)$。我们注意到,以上的计算过程中隐含地假设样本中最有效的决策单位的效率值为

① 为了表述方便,将常数项 α 从 x_{it} 中分出。

1，即它是完全有效的。当样本中的决策单位数量比较多时，这个假设是合理的，因为当 $N \to \infty$ 时，样本中最有效率的决策单位即为完全有效的。组内估计量的优势在于，对效率项（u_i）和随机误差项（v_{it}）不作任何分布假设，并且允许投入（x_{it}）和效率项（u_i）之间存在相关性。然而，固定效应的估计方法将不随时间变化的不可观察的异质性（α_i）包含在 u_i 中，使得效率的估计值是效率真实值与 α_i 的加总，但是 α_i 可能与效率的真实值没有任何关系。

四、效率可随时间变化的 SFA 模型

当面板数据的时间跨度（T）较长时，效率不随时间变化的假设将受到广泛的质疑。正如 Schmidt（1985）①所言，效率项（u_{it}）不能随时间变化是当时 SFA 模型最大的限制，研究者应该致力于允许 u_{it} 随时间改变。

Cornwell 等（1990）[7]首次在面板数据中允许效率项（u_{it}）随时间变化。他们的方法是在模型：

$$y_{it} = \alpha + x'_{it}\beta - u_{it} + v_{it} = x'_{it}\beta + \alpha_{it} + v_{it} \quad (13)$$

式（13）中，令 $\alpha_{it} = \theta_{i1} + \theta_{i2}t + \theta_{i3}t^2$，即 u_{it} 随时间变化的趋势是一个二次函数。此时模型变为：

$$y_{it} = x'_{it}\beta + \theta_{i1} + \theta_{i2}t + \theta_{i3}t^2 + v_{it} = x'_{it}\beta + w'_{it}\delta_i + v_{it} \quad (14)$$

其中，$w_{it} = (1, t, t^2)'$，$\delta_i = (\theta_{i1}, \theta_{i2}, \theta_{i3})'$。令 $w_i = (w_{i1}, w_{i2}, \cdots, w_{iT})'$，$Q = \mathrm{diag}(w_i)$ $(i = 1, 2, \cdots, N)$ 以及 $P_Q = Q(Q'Q)^{-1}Q'$，$M_Q = I - P_Q$，得到模型的矩阵形式 $Y = X\beta + Q\delta + V$，从而得到系数估计 $\hat{\beta} = (X'M_QX)^{-1}(X'M_QY)$。进一步地，对于决策单位 i，用 $y_{it} - x'_{it}\hat{\beta}$ 作为被解释变量，对常数项、时间（t）和时间的二次项（t^2）做最小二乘回归得到 δ_i 的估计 $\hat{\delta}_i$，进而有 $\hat{\alpha}_{it} = w'_{it}\hat{\delta}_i$。令 $\hat{\alpha}_t = max_i(\hat{\alpha}_{it})$，则 $\hat{u}_{it} = \hat{\alpha}_i - \hat{\alpha}_{it}$，效率的估计为 $\exp(-\hat{u}_{it})$。可以发现，通过对效率的变化趋势做具体的参数假设，Cornwell 等（1990）允许了 u_{it} 随时间变化。值得注意的是，在 Cornwell 等（1990）的模型中，不需要对 u_{it} 和 v_{it} 的分布做出假设，并且允许 u_{it} 和 x_{it} 之间存在相关性。不过，$\hat{u}_{it}$ 的一致性依赖于两个条件：（a）$T \to \infty$；（b）效率项的变化趋势是二次函数形式。

与 Cornwell 等（1990）不同，Lee 和 Schmidt（1993）[8]假设：

$$y_{it} = \alpha + x'_{it}\beta - u_{it} + v_{it} = x'_{it}\beta + \alpha_{it} + v_{it} = x'_{it}\beta + \theta_t\delta_i + v_{it} \quad (15)$$

其中，θ_t 可以是任何形式的参数。令 $\xi = (1, \theta_2, \cdots, \theta_T)'$，为了模型可以识别，将 θ_1 标准化为 1，令 $P_\xi = \xi(\xi'\xi)^{-1}\xi'$，$M_\xi = I_T - P_\xi$，可以得到系数估计：

$$\hat{\beta} = \left(\sum_{i=1}^{N} X'_iM_\xi X_i\right)^{-1}\left(\sum_{i=1}^{N} X'_iM_\xi Y_i\right) \quad (16)$$

其中，$X_i = (x_{i1}, x_{i2}, \cdots, x_{iT})'$，$Y_i = (y_{i1}, y_{i2}, \cdots, y_{iT})'$。如果 $\theta_2 = \theta_3 = \cdots = \theta_T = 1$，那么，上述模型简化为标准的面板数据模型。在 $\hat{\beta}$ 的表达式中，ξ 是未知的，Lee 和 Schmidt（1993）证明了 $\hat{\xi}$ 是 $\sum_i (y_i - x_i\hat{\beta})(y_i - x_i\hat{\beta})'$ 最大的特征根对应的特征向量。运用迭代的方法，可以得到 β 和 ξ 的一致估计。Lee 和 Schmidt（1993）建议用固定效应模型中 β 的估计值作为初值，进行迭代以减少迭代的次数。此外，通过最小化损失函数：

$$\sum_{i=1}^{N} (Y_i - X_i\beta - \xi\delta_i)'(Y_i - X_i\beta - \xi\delta_i) \quad (17)$$

可以得到 $\delta_i = \xi'(y_i - x_i\beta)/\xi'\xi$，代入 β 和 ξ 的估计值，可得 δ_i 的估计值 $\hat{\delta}_i$，进而有 $\hat{\alpha}_{it} = \hat{\theta}_t\hat{\delta}_i$。同样地，令 $\hat{\alpha}_t = \max_i(\hat{\alpha}_{it})$，则 $\hat{u}_{it} = \hat{\alpha}_i - \hat{\alpha}_{it}$，效率的估计为 $\exp(-\hat{u}_{it})$。总的来说，与 Cornwell 等（1990）相比，Lee 和 Schmidt（1993）对于效率项（u_{it}）的假设更加一般化，同时允许 u_{it} 和 x_{it} 之间存在相关性。

与 Cornwell 等（1990），Lee 和 Schmidt（1993）没有对 u_{it} 和 v_{it} 的分布做出假设不同，Kumbhakar（1990）[9]在特定的分布假设下，提出了效率可随时间变化的模型。具体讲，对于模型：

$$y_{it} = x'_{it}\beta + \varepsilon_{it} = x'_{it}\beta - u_{it} + v_{it} \quad (18)$$

假设：$u_{it} = \gamma_t u_i$，u_i 服从半正态分布 $N^+(0, \sigma_u^2)$，v_{it} 服从正态分布 $N(0, \sigma_v^2)$，u_i 和 v_{it} 之间相互独立，以及 u_i、v_{it} 与 x_{it} 之间不相关。Kumbhakar（1990）进一步假设：

$$\gamma_t = 1/(1 + \exp(bt + ct^2)) \quad (19)$$

这样的假设不仅能够保证 $\gamma_t \geqslant 0$，从而有 $u_{it} \geqslant 0$，而且根据 b 和 c 的符号，γ_t 可以呈现递增（递减）

① Schmidt P. Frontier Production Functions [J]. Econometric Reviews，1985，4(2)：289－328.

或者凹性(凸性)的特征,保证了效率随时间变化的灵活性。当 $b = c = 0$ 时,$u_{it} = u_i$,即效率不随时间变化。根据:

$$h(\varepsilon_i) = \int_0^\infty \prod_{t=1}^{T} g(\varepsilon_{it} + u_{it}) f(u_i)\, du = \int_0^\infty \prod_{t=1}^{T} g(\varepsilon_{it} + \gamma_t u_i) f(u_i)\, du \quad (20)$$

可以得到相应的似然函数 $L = \prod_{i=1}^{N} h(\varepsilon_i)$,其中,$g(\cdot)$ 和 $f(\cdot)$ 分别表示 v_{it} 和 u_i 的概率密度函数。运用极大似然估计得到参数 β、σ_u^2、σ_v^2、b 和 c 的估计。借鉴 Jondrow 等(1982)的方法,Kumbhakar(1990)得到决策单位 i 在时间 t 的效率:

$$E(\exp(-u_{it}) \mid \varepsilon_i) = E(\exp(-\gamma_t u_i \mid \varepsilon_i)) \quad (21)$$

此外,可以用 LR 统计量对原假设 $H_0: b = c = 0$ 进行检验,判断数据是否支持效率不随时间变化的原假设。

Battese 和 Coelli(1992)[10] 将 Kumbhakar(1990)模型中 u_i 服从半正态分布的假设扩展为从 0 处截断的正态分布,但是他们假设:

$$\gamma_t = \exp(-\eta(t - T)) \quad (22)$$

即假定效率在样本期间内只能随时间单调递增、递减或者不变。我们认为,此假设是十分严格的。如果效率在样本期间内随时间出现先增后减或者先减后增的情况,使用 Battese 和 Coelli(1992)估计效率将会由于模型误设而导致估计的不一致。在实证领域,国内外不少研究运用 Battese 和 Coelli(1992)的模型估计金融机构的效率,得到效率随时间推移而单调递增(或递减)的结论,这可能仅是由于模型设定造成的,而不是效率真的在不断改善(或恶化)。最后,从假设上看,Kumbhakar(1990),Battese 和 Coelli(1992)可以看作是 Lee 和 Schmidt(1993)中 $\theta_t = 1/(1 + \exp(bt + ct^2))$ 或者 $\theta_t = \exp(-\eta(t - T))$ 的特殊形式。

以上介绍的估计方法对应的计量模型均为 $y_{it} = \alpha + x'_{it}\beta - u_{it} + v_{it}$。然而,Greene(2004①,2005②)认为,这样的模型没有考虑决策单位 i 的不可观察的异质性,而忽略异质性的后果就是估计的效率项(u_{it})中包含了单个决策单位的不可观察异质性,而这些异质性可能和真实的效率没有任何关系。因此,Greene(2004,2005)建议考虑如下的计量模型:

$$y_{it} = \alpha_i + x'_{it}\beta + \varepsilon_{it} = \alpha_i + x'_{it}\beta - u_{it} + v_{it} \quad (23)$$

其中,α_i 表示决策单位 i 的不可观察异质性。假设: ⓐ u_{it} 在不同的 i 和 t 之间服从相互独立的半正态分布 $N^+(0,\sigma_u^2)$,ⓑ v_{it} 服从相互独立的正态分布 $N(0,\sigma_v^2)$,ⓒ u_i 和 v_{it} 之间相互独立,ⓓ u_i、v_{it} 与 x_{it} 之间不相关,可以得到 ε_{it} 的概率密度函数:

$$f(\varepsilon_{it}) = \int_0^\infty g(\varepsilon_{it} + u_{it}) f(u_{it})\, du = \frac{2}{\sigma}\varphi\left(\frac{\varepsilon_{it}}{\sigma}\right)\Phi\left(\frac{-\lambda\varepsilon_{it}}{\sigma}\right) \quad (24)$$

式(24)中,$\varepsilon_{it} = y_{it} - \alpha_i + x'_{it}\beta$,$\sigma^2 = \sigma_u^2 + \sigma_v^2$,$\lambda = \frac{\sigma_u}{\sigma_v}$,相应的似然函数为 $L = \prod_{i=1}^{N}\prod_{t=1}^{T} f(\varepsilon_{it})$。运用极大似然估计,Greene(2004,2005)得到参数 β、α_i、σ^2、λ 的估计值,并采用 Jondrow 等(1982)的方法得到效率的估计值。需要注意的是,以上的极大似然估计由于需要同时估计决策单位的异质性(α_i),因此存在冗余参数的问题。Hsiao(1996)③证明了在 Logit 模型中,当 $T = 2$ 时,决策单位异质性(α_i)的存在将导致极大似然估计得到的参数 $\hat{\beta}$ 是不一致的; 然而,Heckman(1981)④对 Probit 模型进行蒙特卡罗模拟后发现,尽管存在 α_i,但是当 $T = 8$、$N = 100$ 时,参数 $\hat{\beta}$ 的偏误已经趋近于 0。正如 Greene(2005)所言,还没有文献讨论过当被解释变量是连续变量时,异质性(α_i)的存在对于参数 $\hat{\beta}$ 一致性的影响以及对效率估计结果的影响。Greene(2005)运用蒙特卡罗模拟发现,当 $N = 500$、$T = 5$ 时,决策单位异质性(α_i)的存在不会对 β 估计的一致性产生影响,但是对 σ_u^2 的估计会出现偏误,从而导致 u_{it} 的估计值与真实值之间出现了 0.05 的偏差(设定 u_{it} 的真实值为 0.25)。针对是否应该在模型中体现不可观察的异质性,学者们也进行了广泛的讨论。争论

---

① Greene W H. Distinguishing between heterogeneity and inefficiency: Stochastic frontier analysis of the world health organization's panel data on national health care systems [J]. Health Economics, 2004, 13(10): 959-980.

② Greene W H. Fixed and random effects in stochastic frontier models [J]. Journal of Productive Analysis, 2005, 23(1): 7-32.

③ Hsiao C. Analysis of panel data [M]. Cambridge: Cambridge University, 1996.

④ Heckman J. The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process, In: Manski, C. and D. McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications [M], Cambridge: MIT Press, 1981.

的关键在于，将效率视为随时间变化、不可观察的异质性设为不随时间变化过于绝对。因此，上述问题并没有得到一致的结论。

Wang 和 Ho(2010)[11] 通过对 u_{it} 做进一步假设，解决了 Greene(2004，2005) 中存在的冗余参数的问题。在 Wang 和 Ho(2010) 的模型中，假设 $u_{it} = h_{it}u_i$，$h_{it} = f(z_{it}\delta)$，u_i 服从从 0 处截断的正态分布 $N^+(\mu,\sigma_u^2)$。对于计量模型：

$$y_{it} = \alpha_i + x'_{it}\beta + \varepsilon_{it} = \alpha_i + x'_{it}\beta - u_{it} + v_{it} \quad (25)$$

首先进行一阶差分，将决策单位的异质性 (α_i) 消去，得到：

$$\Delta y_{it} = \Delta x'_{it}\beta + \Delta\varepsilon_{it} = \Delta x'_{it}\beta - \Delta u_{it} + \Delta v_{it} \quad (26)$$

其中，$\Delta u_{it} = \Delta h_{it}u_i$。然后，对差分后的计量模型做极大似然估计。从这种做法的思想中可以看出，Wang 和 Ho(2010) 之所以假设 $u_{it} = h_{it}u_i$，是因为差分后 u_i 没有发生变化，因此可以得到 $\Delta u_{it} = \Delta h_{it}u_i$ 的概率密度函数，进而得到 $\Delta\varepsilon_{it}$ 的概率密度函数。如果象 Greene(2004，2005) 一样假设 u_{it} 在不同的 i 和 t 之间服从相互独立的半正态分布 $N^+(0,\sigma_u^2)$，就无法得到 Δu_{it} 的分布函数的具体表达形式。由于决策单位的异质性 α_i 已经从模型中消去，因此不存在冗余参数的问题。参考 Jondrow 等(1982) 的方法，Wang 和 Ho(2010) 建议用 $E(\exp(-u_i t \mid \Delta\varepsilon_i))$ 得到不同决策单位 i 在时间 t 的效率，其中，$\Delta\varepsilon_i = (\Delta\varepsilon_{i1},\Delta\varepsilon_{i2},\cdots,\Delta\varepsilon_{iT})'$。通过蒙特卡罗模拟，Wang 和 Ho(2010) 发现，与 Greene(2004，2005) 相比，先差分再做极大似然估计的方法得到的效率的估计更加精确，比如，当 N = 100、T = 50 时，两种方法得到的效率的估计值与真实值之间的相关系数分别为 0.871 和 0.711。

五、研究效率影响因素的 SFA 模型

在估计完效率值后，人们还经常关心效率的影响因素有哪些。对此，通常是采用“两步法”，即首先忽略效率的影响因素，采用极大似然方法估计效率值，然后再将效率值作为被解释变量，对效率的影响因素 $(z_{it}: L \times 1)$ 进行分析。“两步法”可以根据某两个影响因素进行分组统计，但更多的是进行多个影响因素的回归分析，具体做法主要有以下四种。一是，直接对效率值进行最小二乘回归，这包括了不少采用 Granger 因果检验的研究。这种做法没有考虑效率值的取值范围是 [0,1] 的特点。二是，对效率值进行 Logit 变化，再对变换后的效率值(取值范围为 $(-\infty,+\infty)$) 进行最小二乘估计。不过，变换后的效率值失去了明确的经济含义①。三是，对效率值进行 Tobit 回归。不过，效率只是被“定义”为实际值与最优值的比值，并不能找到效率背后的潜变量，所以效率值并不属于删失数据(censored data)，而是属于比例数据。四是，采用 Papke 和 Wooldridge(1996)[12] 的比例因变量模型②。该方法采用拟极大似然方法估计“第二步”回归中的参数，只要条件期望的设定是正确的，参数的估计就是一致的，避免了极大似然估计由于分布的误设带来的偏误。

“两步法”被广泛用于研究效率的影响因素，然而，“两步法”由于在第一步运用极大似然估计效率时，忽略了投入 (x_{it}) 和效率的影响因素 (z_{it}) 之间的相关性，可能导致参数估计结果存在偏误(Wang 和 Schmidt，2002[13])。对此，一些学者研究了如何在估计出参数 (β) 的同时，也估计出效率的影响因素，这被称为“一步法”。

Battese 和 Coelli(1995)[14] 在面板数据中首次建立了“一步法”的模型。具体地讲，他们假设 u_{it} 服从在不同的 i 和 t 之间相互独立的、从 0 处截断的正态分布 $N^+(\mu_{it},\sigma_u^2)$，其中，$\mu_{it} = z_{it}\delta$，z_{it} 表示 L 个影响效率的因素；其他条件与 Battese 和 Coelli(1992) 相同。根据以上假设，可以得到 ε_{it} 的概率密度函数：

$$\begin{aligned} f(\varepsilon_{it}) &= \int_0^\infty g(\varepsilon_{it} + u_{it}) f(u_{it})\, du \\ &= \int_0^\infty \frac{\exp\left(-\frac{1}{2}(\varepsilon_{it} + u_{it})^2/\sigma_v^2\right)\exp\left((u_{it} - z'_{it}\delta)/\sigma_u^2\right)}{2\pi\sigma_u\sigma_v\Phi(z_{it}\delta/\sigma_u)} du \end{aligned} \quad (27)$$

以及相应的似然函数 $L = \prod_{i=1}^{N}\prod_{t=1}^{T} f(\varepsilon_{it})$。运用极大似然估计可以得到效率的影响因素的参数 (δ) 的一致估计。

Wang(2002)[15] 在 Battese 和 Coelli(1995) 的基础上，进一步考虑了方差 σ_u^2 和 σ_v^2 的异质性。具体地讲，他们假设 $\sigma_{u_{it}}^2 = \exp(z'_{it}\gamma)$，$\sigma_{v_{it}}^2 = \exp(z'_{it}\eta)$。

① 对于表示一类事项发生可能性的“概率值”(p) 进行 Logit 变化，得到“对数机率比”$(\log(p/(1-p)))$，是有经济含义的。

② Papke 和 Wooldridge(1996) 提出该方法并首先用于研究美国雇员对 401(k) 养老金计划的参与率。

Wang(2002)的模型较之 Battese 和 Coelli(1995)的模型的优势在于考虑了方差异质性,并且允许效率的影响因素对效率的"边际效应"存在非单调的变化,边际效应的正负号也可以随影响因素的值发生变化。具体地讲,由于:

$$E(u_{it}) = \mu_{it} + \sigma_{u_{it}}^{2}\frac{\phi(\Lambda)}{\Phi(\Lambda)} \tag{28}$$

其中,$\Lambda = \frac{u_{it}}{\sigma_{it}}$,则 z_{it} 中第 k 个因素($z_{it}(k)$)对 $E(u_{it})$ 的边际效应如下:

$$\frac{\partial E(u_{it})}{\partial z_{it}(k)} = \delta(k)\left(1 - \Lambda\frac{\phi(\Lambda)}{\Phi(\Lambda)} - \left[\frac{\phi(\Lambda)}{\Phi(\Lambda)}\right]^2\right) + \gamma(k)\frac{\sigma_{u_{it}}^{2}}{2}\left((1+\Lambda^2)\frac{\phi(\Lambda)}{\Phi(\Lambda)} + \Lambda\left(\frac{\phi(\Lambda)}{\Phi(\Lambda)}\right)^2\right) \tag{29}$$

从式(29)可以看出,边际效应同时取决于参数 $\delta(k)$、$\gamma(k)$ 以及相应的权重。伴随着不同的 Λ 值,边际效应的正负号可以发生变化。然而,在 Battese 和 Coelli(1995)中,由于 $\sigma_{u_{it}}^{2} = \sigma_{u}^{2}$,即 $\gamma(k) = 0$,将会得到:

$$\frac{\partial E(u_{it})}{\partial z_{it}(k)} = \delta(k)\left(1 - \Lambda\frac{\phi(\Lambda)}{\Phi(\Lambda)} - \left(\frac{\phi(\Lambda)}{\Phi(\Lambda)}\right)^2\right) \tag{30}$$

又因为 $\left(1 - \Lambda\frac{\phi(\Lambda)}{\Phi(\Lambda)} - \left(\frac{\phi(\Lambda)}{\Phi(\Lambda)}\right)^2\right) > 0$,因此 Battese 和 Coelli(1995)的模型中必然会出现 $E(u_{it})$ 对 z_{it} 中第 k 个影响因素 $z_{it}(k)$ 的边际效应与 $\delta(k)$ 的正负号相同。允许边际效应存在非单调的变化具有非常重要的实际意义。比如:考虑体力劳动者的年龄对效率的影响,随着年龄的增长,劳动者的经验越来越多,年龄对效率存在正向影响;但年龄超过一定的范围后,脑力和体力开始下降,年龄对生产效率的影响可能会转化为负。除此之外,企业规模、多元化程度、很多财务指标等对企业效率的边际效应同样可能存在非单调的变化。Battese 和 Coelli(1995)的模型无法体现边际效应的非单调性是一个较严重的缺陷。

六、总结

近年来,面板数据随机前沿分析越来越多地被用于测算各类决策单位的效率。本文在统一的计量框架下,系统地梳理和分析了面板数据 SFA 模型,以期对 SFA 领域的实证研究在方法选择上提供参考建议。根据对效率随时间变化的假设不同,SFA 模型分为两类:效率不随时间变化的模型和效率可随时间变化的模型;在每一类别中,基于是否对效率项(u_{it})的分布函数做出假设,又进一步分为有分布假设的模型和无分布假设的模型。同时,本文对研究效率的影响因素的模型进行了梳理。

SFA 方法在研究国内外各类机构的效率问题上有着大量应用,取得了很多成果。不过,这些研究主要依赖几种假设严格的 SFA 模型,一些假设更为灵活的 SFA 模型极少得到运用。针对这种状况,并考虑到较之 DEA 等非参数方法,SFA 等参数方法的劣势主要是面临模型误设风险,因此,采用更灵活的假设是应用 SFA 等参数方法的发展方向,我们提出如下具体建议:

第一,由于效率不随时间变化的假设十分严格,我们建议采用效率可随时间变化的模型估计决策单位的效率值。只有当所有数据的时间跨度很短(比如 $T<3$)或者确信决策单位的效率在样本期间内保持不变时,再考虑采用效率不随时间变化的模型。

第二,在效率不随时间变化的模型中,Pitt 和 Lee(1981)是 Battese 和 Coelli(1988)的一种特殊情形,虽然 Schmidt 和 Sickless(1984)没有对 u_{it} 和 v_{it} 做出分布假设,但效率估计的一致性估计依赖于 $T \to \infty$。根据中心极限定理,在大样本中,相互独立的决策单位对应的随机变量的加总趋近于正态分布,因此,我们建议同时运用 Battese 和 Coelli(1988),Schmidt 和 Sickless(1984)的模型估计效率。与此同时,为了加强效率估计结果的稳健性,建议根据效率的大小对各个决策单位进行排序,观察和判断效率值和排序是否受到了估计方法的较大影响。

第三,在效率可随时间变化的模型中,Battese 和 Coelli(1992)关于效率只能单调递增或递减的假设十分严格,存在较大的模型误设的风险,不建议在实证应用中采用;Cornwell 等(1990)对于效率随时间变化的趋势是二次函数的假设也很严格,而且此假设可能导致样本内期初或者期末的效率的估计存在较大的偏误。我们建议今后的研究中使用 Kumbhakar(1990),Lee 和 Schmidt(1993)的模型估计效率。

第四,如果要进一步研究效率的影响因素,"一步法"和"两步法"各有优劣。在"一步法"模型中,Wang(2002)同时考虑了效率项(u_{it})的均值和方差的异质性,并且允许边际效应存在非单调的变化,是

更为灵活的方法。但是，"一步法"需要对 u_{it} 和 v_{it} 的分布做出先验的假设。在"两步法"模型中，我们建议采用 Lee 和 Schmidt(1993) 估计效率，然后运用 Papke 和 Wooldridge(1996) 估计效率的影响因素。这种组合的优势在于前者不需要假设 u_{it} 和 v_{it} 的分布，后者采用拟极大似然估计从而对不同的分布假设保持稳健。

最后，我们发现，Battese 和 Coelli(1988，1992，1995) 的模型得到了很广泛的使用，这在很大程度上归因于 Coelli 编写了研究效率的 Frontier 软件，而 Frontier 软件采用的主要就是 Battese 和 Coelli(1988，1992，1995) 的模型。因此，相关软件的开发者如果能将 Kumbhakar(1990)、Lee 和 Schmidt(1993)、Wang 和 Ho(2010)、Wang(2002) 等模型编入应用程序中，对后续实证研究是大有裨益的。

参考文献

[1] Koopmas T. An analysis of production as an efficient combination of activities. In: Activity analysis of production and allocation, Vol. 13 of Cowles Commission for Research in Economics [M]. New York: John Wiley and Sons Press, 1951.

[2] Debreu G. The coefficient of resource utilization [J]. Econometrica, 1951, 19(3): 273 - 292.

[3] Farrell M J. The measurement of productive efficiency [J]. Journal of the Royal Statistical Society. Series A (General), 1957, 120(3): 253 - 290.

[4] Pitt M, Lee L. The measurement and sources of technical inefficiency in the Indonesian weaving industry [J]. Journal of Development Economics, 1981, 9(1): 43 - 64.

[5] Jondrow J, Lowell C, Materov I, et al. On the estimation of technical efficiency in the stochastic production function model [J]. Journal of Econometrics, 1982, 19(2/3): 233 - 238.

[6] Schmidt P, Sickles R. Production frontiers and panel data [J]. Journal of Business Economics and Statistics, 1984, 2(4): 367 - 374.

[7] Cornwell C, Schmidt P, Sickles R. Production frontiers with cross-sectional and time-series variation in efficiency levels [J]. Journal of Econometrics, 1990, 46(1/2): 185 - 200.

[8] Lee Y, Schmidt P. A production frontier model with flexible temporal variation in technical inefficiency. In: Fried, H., C. Lowell and P. Schmidt (eds.), The Measurement of Productive Efficiency: Techniques and Applications [M]. Oxford: Oxford University Press, 1993.

[9] Kumbhakar S. Production frontiers, panel data and time-varying technical inefficiency [J]. Journal of Econometrics, 1990, 46(1/2): 201 - 212.

[10] Battese G, Coelli T. Frontier production function, technical efficiency and panel data: with application to paddy farmers in India [J]. Journal of Productivity Analysis, 1992, 3(1/2): 153 - 159.

[11] Wang H, Ho C. Estimating fixed-effect panel stochastic frontier models by model transformation [J]. Journal of Econometrics, 2010, 157(2): 286 - 296.

[12] Papke L E, Wooldridge J M. Econometric methods for fractional response variables with an application to 401(k) plan participation rates [J]. Journal of Applied Econometrics, 1996, 11(6): 619 - 632.

[13] Wang H, Schmidt P. One-step and Two-step estimation of the effects of exogenous variables on technical efficiency levels [J]. Journal of Productivity Analysis, 2002, 18(2): 129 - 144.

[14] Battese G, Coelli T. A model for technical inefficiency effects in a stochastic frontier production function for panel data [J]. Empirical Economics, 1995, 20(2): 325 - 322.

[15] Wang H. Heteroscedasticity and non-monotonicity efficiency effects of a stochastic frontier model [J]. Journal of Productivity Analysis, 2002, 18(3): 241 - 253.

作者简介

边文龙，男，2016 年毕业于北京大学国家发展研究院，获经济学博士学位，现为广东外语外贸大学金融学院讲师。研究方向为计量经济学及其应用、银行和保险。

王向楠，男，2012 年毕业于西南财经大学保险学院，获经济学博士学位，现为中国社会科学院金融研究所助理研究员。研究方向为金融与保险。

(责任编辑：倪立行)