行政院國家科學委員會專題研究計畫　成果報告

隨機邊界模型檢定的小樣本問題

計畫類別：個別型計畫
計畫編號：NSC92-2415-H-032-011-
執行期間：92 年 08 月 01 日至 93 年 07 月 31 日
執行單位：淡江大學經濟系(所)

計畫主持人：陳怡宜

計畫參與人員：陳怡宜

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 93 年 12 月 20 日

Small Sample Properties of Hypothesis Testings and Confidence Intervals of a Stochastic Frontier Model

1. Introduction

The contribution of stochastic frontier models to the empirical production literature lies in the accommodation of production inefficiencies in the empirical models. A stylized stochastic frontier model in a panel data setting may be written as (Aigner, Lovell, and Schmidt 1977)

$$\ln y_{it} = \ln f(\boldsymbol{x}_{it}; \beta) + v_{it} - u_i, \tag{1}$$

$$e_{it} \equiv v_{it} - u_i, \tag{2}$$

where y_{it} is the output of producer i, $\boldsymbol{x}_{it}$ is a vector of inputs, v_{it} is the zero-mean statistical error, and $u_i > 0$ is the producer-specific inefficiency. The u_i is often assumed to be a random variable following a specific distribution such as the half-normal distribution. In this specification, $f(\cdot) + v_{it}$ is the stochastic production frontier, and u_i is the deviation from the frontier due to inefficiency. The model is complete with the following distribution assumptions:

$$v_{it} \sim N(0, \sigma^2), \tag{3}$$

$$u_i \sim N^+(0, \sigma_u^2), \tag{4}$$

where $N^+(\cdot)$ denotes non-negative truncation of the normal distribution.

It is clear that the term u_i is the only thing that distinguishes (1) from a standard production function. Therefore, the empirical justification of a stochastic frontier model lies in a large part in the statistical importance of u_i. If u_i is important and significant, omitting the term would mean losses of important information about the production behavior. If, on the other hand, u_i does not have the needed statistical justification, the more simple classical production function should be used in lieu of the elaborated stochastic frontier model.

The literature has developed statistics to test the significance of u_i. Coelli (1995) and Lee (1993) develop a one-sided likelihood ration (LR) test to test the overall significance of u_i. The distribution of the statistic is asymptotically equal to a mixture of c^2 distribution $0.5\, c_0^2 + 0.5\, c_1^2$, where c_0^2 has the unit mass at zero. The Monte Carlo evidence of Coelli (1995), however, showed that the test has an unsatisfactory power in finite samples.

In addition to the test of the *overall* significance of the production inefficiency, the test of the *individual* producer's inefficiency is also important for studies of production efficiency. The idea is that, in addition to knowing whether the sample exhibits aggregate inefficiency, it may also of great interests to know *which* producer is more efficient than the others. Identifying the more efficient producers or knowing the efficiency differences between firms and groups can have important policy implications. For instances, properties of the efficient producers can be studied and their operation characteristics be promoted. Also, government subsidy or other incentive-compatible programs can be better designed and targeted.

To test the individual importance of u_i, Horrace and Schmidt (1996) proposes constructing the confidence interval (CI) of $E(u_i \mid \hat{e}_{it})$. The authors show that the CI is easy to construct as long as we recognize that the distribution of $E(u_i \mid \hat{e}_{it})$ follows

$$E(u_i \mid \hat{e}_{it}) \sim N^{+}(m_{it}^{*}, s_{*}^{2}), \quad (5)$$

where

$$m_{it}^{*} = -e_{it}\, s_{u}^{2}\, (s_{u}^{2} + \sigma^2)^{-1}, \quad (6)$$

$$s_{*}^{2} = s_{u}^{2}\, \sigma^2\, (s_{u}^{2} + \sigma^2)^{-1}. \quad (7)$$

The significance of u_i is viewed by whether the associated CI encompasses the value of zero, and the question of whether $u_i = u_j$, $i \neq j$, is tested by whether the CI of $E(u_i|\hat{e}_{it})$ contains the value of $E(u_j|\hat{e}_{jt})$.

Although there is no formal study on the accuracy of the CI constructed above in finite samples, empirical examples provided by Horrace and Schmidt (1996) indicate that the constructed CI seems to be *too wide*, in the sense that $E(u_i|\hat{e}_{it})$ usually cannot be told apart from $E(u_j|\hat{e}_{jt})$ based on the CIs.

In this study, we conduct a detailed Monte Carlo analysis on the small sample performance of the statistics concerning $\hat{u}_i$, with a particular emphasis put on the statistic of Horrace and Schmidt (1996). In addition, we also consider an alternative testing procedure based on the bootstrap approach.

2. Design of the Monte Carlo Analysis

The Monte Carlo model is the follows.

$$y_{it} = b_0 + b_1 x_{it}, + v_{it} - u_{i}, \quad (8)$$

$$x_{it} \sim N(0, 1), \quad (9)$$

$$v_{it} \sim N(0, \sigma^2), \quad (10)$$
$$u_{it} \sim N(0, s_u^2), \quad (11)$$

The base case (Case 1) has the following specifications.

Case 1: $b_0=0.5$; $b_1=0.5$; $\sigma^2 = 1$; $s_u^2 = 2$; $N=50$; $T=2$.

We first generate a set of values of x_{it}, u_i, v_{it}, and y_{it} to be the "true" data for the subsequent use in bootstrapping. The asymptotic-based test statistics of Horrace and Schmidt (1996) are calculated according to (5), (6), and (7), and nonparametric bootstrap results, for both percentile and the bias-corrected methods, are also recorded. We then compare the power of the two statistics.

Alternative cases are designed as the follows.

Case 2: $b_0=0.5$; $b_1=0.5$; $\sigma^2 = 1$; $s_u^2 = 3$; $N=50$; $T=2$.
Case 3: $b_0=0.5$; $b_1=0.5$; $\sigma^2 = 1$; $s_u^2 = 2$; $N=50$; $T=5$.

Table 1 reports the power of the statistics for a 5% statistical test on the estimates of $E(u_j|\hat{e}_{jt})$.

Table 1: Power of the Test Statistics

	HS96	BT_p	BT_bc
Case 1	0.029	0.006	0.043
Case 2	0.011	0.014	0.026
Case 3	0.008	0.020	0.002

Note: HS96: Horrace and Schmidt (1996) statistic; BT_p: percentile bootstrap statistic; BT_bc: bias-corrected bootstrap statistic.

3. Discussions

The results show that the small sample performance of either the asymptotic method or the bootstrap method are very disappointing. The finding of under-size of the asymptotic method is consistent with the empirical finding of Horrace and Schmidt (1996) who documented that the constructed confidence intervals are too wide. On the other hand, the under-size of the bootstrap method may not be surprising as well. For the percentile method, reading off the α and 1-α percentile interval endpoints of the bootstrap distribution amounts to assuming that the statistic in study is unbiased and its distribution scale invariant. This is clearly not the case for a study of the efficiency distribution. As well, the percentile-*t* interval require reliable estimates of scale. However, given that the location and mean are both complicated functions of the model's parameters which have small sample bias as shown by Coelli (1995), a reliable estimates of the scale is unlikely.

The results of this study raises doubt on the empirical application of the hypothesis testing regarding the significance of the individual inefficiency level. Clearly, devising a better test statistic will be important for the literature.

Reference:

Aigner, Dennis, C. A. Knox Lovell, and Peter Schmidt. (1977) "Formulation and Estimation of Stochastic Frontier Production Function Models", *Journal of Econometrics* 6, pp. 21-37

Coelli, Tim. (1995) "Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis", *Journal of Productivity Analysis* 6, pp. 247-68

Gourieroux, Christian, Alberto Holly, and Alain Monfort. (1982) "Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters", *Econometrica* 50, pp. 63-80

Horrace, William C., and Peter Schmidt. (1996) "Confidence Statements for Efficiency Estimates from Stochastic Frontier Models", *Journal of Productivity Analysis* 7, pp. 257-82

Jondrow, Lovell, Materov, and Schmidt. (1982) "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model", *Journal of Econometrics* 19, pp. 233-38

Lee, Lung Fei. (1993) "Asymptotic Distribution of the Maximum Likelihood Estimator for a Stochastic Frontier Function Model with a Singular Information Matrix", *Econometric Theory* 9, pp. 413-430