No. 1063

Empirical Evidences for the Skewed Distribution of Inefficiency

by
Yoshihiko Tsukuda and Tatsuyoshi Miyakoshi

November 2003

Empirical Evidences for the Skewed Distribution of Inefficiency

Yoshihiko Tsukuda
Tohoku University, Graduate School of Economics,
Kawauchi, Aoba-ku, Sendai, 980-8576, JAPAN.
and
Tatsuyoshi Miyakoshi*
University of Tsukuba, Institute of Policy and Planning Sciences,
Tsukuba, Ibaraki, 305-8573, JAPAN.
November 7, 2003

Abstract:

Most of the previous researches have ignored the fact that the distribution of technical inefficiency is always asymmetric and have routinely used the expectation of the conditional distribution for predicting the technical inefficiency in stochastic frontier production models. The median adequately exhibits the representative value of the distribution and is possibly quite different from the expected value. An extensive examination of the previous empirical studies in fact indicates that the distributions are heavily skewed for some of the data sets. The paper cautions that careless use of the conditional expectation may mislead the evaluation of technical inefficiency in practice.

Keywords: Stochastic frontier; Technical inefficiency; Prediction; Median
JEL Classification Numbers: C23; D24
A running title: Empirical Evidences for the Skewed Distribution of Inefficiency

---

* The research by the second author was financially supported by Postal Life Insurance of Japan in 2002.
Corresponding author: Tatsuyoshi Miyakoshi, Institute of Policy and Planning Sciences, University of Tsukuba, Tsukuba, Ibaraki, 305-8573, JAPAN.
Tel: + 81 – 298 – 53 - 5168; Fax: + 81 - 298 - 55 - 3849.
E-mail address: miyakosi@sk.tsukuba.ac.jp

1. Introduction

We consider a stochastic frontier production function by Battese and Coelli (1995) including most of the previous functions in this field:

$$Y_{it} = \exp(X_{it}\beta + \varepsilon_{it}) \ , \qquad (1)$$
$$\varepsilon_{it} = V_{it} - U_{it} \ , \qquad i = 1,\dots,N \ ; \ t = 1,\dots,T,$$

where Y_{it} denotes the output for the *i-th* firm at the *t-th* time period; X_{it} is a 1×k vector of k input variables; β is a k×1 vector of k unknown parameters; N is the number of firms and T is the number of time periods. Random variables V_{it} are assumed to have iid $N(0,\sigma_V^2)$ distribution, non-negative random variables U_{it} are assumed to have iid $N(\mu_{it},\sigma_U^2)$ distribution that is truncated at zero, and $\mu_{it} = Z_{it}\delta$ where Z_{it} *is* a 1×p vector of variables which influence the inefficiency of a firm and δ is a p×1 vector of unknown parameters. The V_{it} are assumed to be independent of U_{it} . The U_{it} accounts for technical inefficiency in production for the *i-th* firm at the *t-th* time. Battese and Coelli (1988, p.389) define the rate of technical efficiency of production for the *i-th* firm at the *t-th* time as a ratio of its mean production to the corresponding mean with $U_{it} = 0$, i.e. $TE_{it} = E(Y_{it} \mid U_{it}, X_{it}) / E(Y_{it} \mid U_{it} = 0, X_{it})$. Alternatively the rate of technical inefficiency is defined by $TIE_{it} = 1 - TE_{it}$. It becomes

$$TIE_{it} = 1 - \exp(-U_{it}), \qquad (2)$$

which is a random variable taking the values between zero and one.

The measurement of technical inefficiency is an important issue for the stochastic frontier model. A commonly used predictor for technical inefficiency of production for the *i-th* firm at the *t-th* time conditional on the value of ε_{it} is given by its conditional expectation:

$$ETIE(\mu_{it} \mid \varepsilon_{it}) \equiv E(TIE_{it} \mid \varepsilon_{it})$$

$$= 1 - \{\exp(-\mu^*_{it} + \sigma^{*2}/2)\}\{\Phi((\mu^*_{it}/\sigma^*) - \sigma^*)/\Phi(\mu^*_{it}/\sigma^*)\}, \quad (3)$$

where $\mu^*_{it} \equiv (\sigma_V^2 \mu_{it} - \sigma_U^2 \varepsilon_{it})/(\sigma_V^2 + \sigma_U^2)$, $\sigma^* \equiv \sqrt{\sigma_U^2 \sigma_V^2 /(\sigma_V^2 + \sigma_U^2)}$, and $\Phi(\cdot)$ denotes the standard normal distribution function. We note that μ_{it}^* is a linear function of μ_{it}. Equation (3) was derived by Battese and Coelli (1993, p.20). The values of (3) depend on the observations of Y_{it} and X_{it} only through ε_{it}. We are able to estimate ε_{it} by residuals $\hat{\varepsilon}_{it} = Y_{it} - X_{it}\hat{\beta}$ if we have an estimate $\hat{\beta}$.

Tsukuda and Miyakoshi (2003) proposed the median of the conditional distribution for predicting the technical inefficiency motivated by the fact that the U_{it} is always asymmetrically distributed because it is truncated at zero:

$$MTIE(\mu_{it} \mid \varepsilon_{it}) = 1 - \exp\{-\mu_{it}^* + \sigma^* \Phi^{-1}(0.5\Phi(\mu_{it}^*/\sigma^*))\} \quad . \quad (4)$$

The median is identical to the expected value when the distribution is symmetric, but they can be different for the case of asymmetric distribution. In the latter case, the conditional expectation of (3) may not exhibit the representative value of the distribution. Hence, it is not appropriate measure for the technical inefficiency. Tsukuda and Miyakoshi (2003) compared the two predictors numerically as well as graphically for the various sets of key parameters μ_{it}^* and σ^* and showed that the two predictors take substantially different values when the distribution is heavily asymmetric.

The purpose of the paper is to empirically examine the skewness of the distribution for technical inefficiency from the previous studies appeared on recent articles in the major journals in the field. Most of the previous researches have ignored the fact that the distribution of technical inefficiency is always asymmetric and have routinely reported the expected value of the conditional distribution for predicting the technical inefficiency. Our investigation reveals that the distributions of technical inefficiency in some of the previous studies are heavily skewed depending upon the estimated values of key parameters μ_{it}^* and σ^*. The result implies that careless use of the conditional expectation possibly misleads the evaluation of technical inefficiency in practice.

2. Empirical evidences from the previous studies

First, we investigate how the skewness of the distribution determines the values of the ETIE and the MTIE. The skewness S_K is derived as follows:

$$S_k = -\{\exp(3\sigma^{*2})\Phi(\Delta)^2\Phi(\Delta-3\sigma^*) - 3\exp(\sigma^{*2})\Phi(\Delta)\Phi(\Delta-\sigma^*)\Phi(\Delta-2\sigma^*)$$

$$+2\Phi(\Delta-\sigma^*)^3\}\{\exp(\sigma^{*2})\Phi(\Delta)\Phi(\Delta-2\sigma^*) - \Phi(\Delta-\sigma^*)^2\}^{-\frac{3}{2}}, \qquad (5)$$

where $\Delta = \mu_{it}^*/\sigma^*$ (the suffices i and t are omitted from Δ for simplicity of exposition). The derivation of (5) is given in Appendix. Table 1 shows the values of conditional expectation, median and skewness for some specific sets of parameters in the density functions of TIE_{it} conditional on ε_{it}. For example, when $\mu_{it}^* = 1.5$, $\sigma^* = 1.0$, we have $S_k = -1.227$, ETIE = 0.727, MTIE = 0.795, and then the difference between them is 0.068. The two predictors can take different values when the distribution is heavily skewed, while there is no substantial difference when the skewness is small.

Next, we extensively examine the previous empirical studies on the stochastic frontier models, which appeared on the major journals in recent years. We choose 60 cases from 26 articles. The referred articles are Abdulai (2001), Audibert(1997), Apezteguia and Garate(1997), Battese and Coelli(1995), Bhattacharyya, et al (1996), Cullinance, et al(2002), Delorme, et al(1995), Ferrantino and Ferrier(1995), Hattori(2002), Kaynak and Pagan(2003), Mercedes(2000), Morrison, et al(2000), Murillo-Zamorano et al (2001), Pascoe, et al (2002), Puig-Junoy (2001), Seyoum, et al(1998), Shao and Lin(2001), Sharma, et al(1997), Sharma and Leung(2000), Tian(2000), Uri(2001), Wadud and White(2000), Wang(2000), Wilson, et al(2001), Yao, et al (2001). Some articles run the analysis of a model for several data sets within an article. All of the researches cited above ignore the fact that the distribution of technical inefficiency is always asymmetric. Most of them either report the expected values of the conditional distribution in a form of summary statistics such as the maximum, minimum and mean (median, mode) or simply draw the graph of the expected values.

Table2 shows 10 articles out of 26 in the descending order of the estimated value of $\sigma*$. The values of $\sigma*$ vary from 0.621 (Uri(2001)) to 0.130 (Wang(2000)). The values of $\sigma*$ are greater than 0. 3 for the five cases. Table2 gives a rough idea about the practical values of the parameter $\sigma*$. We need the value of μ_{it}^* as well as $\sigma*$ in order to obtain the ETIE, MTIE and skewness of the distribution . (See equations (3), (4) and (5).) Unfortunately, most of the previous studies do not show the value of μ_{it}^* for each i and t, but report only such summary statistics as the maximum, minimum and mean (median, mode) values of the ETIE. However, we can solve the equation (3) for μ_{it}^* given the value of the left hand side and $\sigma*$ since it is monotone increasing in μ_{it}^* (Tsukuda and Miyakoshi (2002)).

The first column in Table 3 denotes the values of the ETIE reported in the first five articles of Table2. The values of the ETIE widely vary from 0.019 to 0.703. In Kaynak and Pagan (2003), for example, the minimum, median and maximum values are 0.019, 0.082 and 0.465 respectively. The second column shows the calculated values of μ_{it}^* from equation (3). The skewness in column 3 is obtained from equation (5) given μ_{it}^* and $\sigma*$. The skewness is scattered from -1.24 to 1.353.

We consider the relationship between the MTIE and the ETIE. The MTIE in column 4 of Table3 is calculated from equation (4). The differences between the ETIE and the MTIE are less than 0.05 for all studies. For example, in Kaynak and Pagan (2003) the difference is 0.02 (= 0.082 - 0.062 = (1.B) - (4.B)) at the median. This value can be expected from the numerical comparison of Table1. The estimates of key parameters ($\mu*$, $\sigma*$) = (-1.825, 0.423) in Kaynak and Pagan (2003) are close to the cell ($\mu*$, $\sigma*$) = (0.0, 0.3) in Table1, in which the difference is 0.018 (= 0.201 - 0.183). The column (5) denotes the relative magnitudes of difference at the median (mean). In Kaynak and Pagan (2003) the relative difference is 24%, which corresponds to 1.353 of skewness. On the other hand, in Cullinance, et al (2002) it is 3% corresponding to 0.248 of skewness.

Another way of looking at the relative magnitude is to remind the fact the mean (mode, median) value of the ETIE is very close to its minimum value in most of the previous studies examined. This fact implies that many values of the ETIE gather around the minimum value. Hence, we could evaluate the difference between the two predictors by using the ratio of the difference to the difference between the mean (mode, median) and minimum of the ETIE as shown in the last column of Table3. In Kaynak and Pagan (2003),the difference between the two predictors is 0.02, but the difference between the median and the minimum of the ETIE is 0.063, hence the ratio is 0.32. From columns (5) and (6), we observe that the differences between the ETIE and the MTIE are not negligible in the sense of the relative magnitudes. The ETIE may mislead the evaluation of technical inefficiency.

3. Conclusion

Tsukuda and Miyakoshi (2003) proposed the median of the conditional distribution for predicting the technical inefficiency in stochastic frontier production models with panel data motivated by the fact that the distribution of technical inefficiency is always asymmetric. An extensive examination of the previous empirical studies indicates that the differences between the ETIE and the MTIE are not negligible in the sense of relative magnitudes, and careless use of the expected technical inefficiency measure may mislead the evaluation of technical inefficiency. The paper empirically supports the use of median instead of the expectation of the conditional distribution.

Appendix: Derivations of (5)

We omit the subscripts i and t in the following. Battese and Coelli (1993, p.20) derived the density of U_{it} conditional on ε_{it} as

$$f_U(U \mid \varepsilon) = \frac{\phi((U-\mu^*)/\sigma^*)}{\sigma^*\Phi(\mu^* \;/\sigma^*)}, \; U \geq 0 \tag{A.1}$$

where $\phi(\cdot)$ stands for the density function of the standard normal distribution. The j-th moment of exp(-U) given ε is derived as

$$\begin{aligned} \eta_j &= E\{e^{-jU}\} \\ &= \int_0^\infty e^{-jU} f_U(u \mid \varepsilon) du, \\ &= \exp\left\{\frac{1}{2}(j\sigma^*)^2 - j\mu^*\right\}\Phi(\Delta)^{-1}\Phi(\Delta - j\sigma^*). \end{aligned} \tag{A.2}$$

Using (A.2), we have the first three central moments of TIE_{ij} ($\equiv$W) as

$$\begin{aligned} m_1 &= E\{W \mid \varepsilon\} \\ &= 1 - \eta_1 \\ &= 1 - \exp\{(1/2)\sigma^{*2} - \mu^*\}\Phi(\Delta)^{-1}\,\Phi(\Delta - \sigma^*) \end{aligned} \tag{A.3}$$

$$\begin{aligned} m_2 &= E\{(W - m_1)^2 \mid \varepsilon\} \\ &= \eta_2 - \eta_1^2 \\ &= \exp\{\sigma^{*2} - 2\mu^*\}\Phi(\Delta)^{-2}\{\exp\{\sigma^{*2}\}\Phi(\Delta)\Phi(\Delta - 2\sigma^*) - [\Phi(\Delta - \sigma^*)]^2\} \end{aligned} \tag{A.4}$$

$$\begin{aligned} m_3 &= E\{(W - m_1)^3 \mid \varepsilon\} \\ &= -\{\eta_3 - 3\eta_1\eta_2 + 2\eta_1^3\} \\ &= -\exp\{(3/2)\sigma^{*2} - 3\mu^*\}\Phi(\Delta)^{-3}\{\exp\{3\sigma^{*2}\}\Phi(\Delta)^2\Phi(\Delta - 3\sigma^*) \\ &\quad - 3\exp\{\sigma^{*2}\}\Phi(\Delta)\Phi(\Delta - \sigma^*)\Phi(\Delta - 2\sigma^*) + 2[\Phi(\Delta - \sigma^*)]^3\} \end{aligned} \tag{A.5}$$

Then, using (A.4) and (A.5), the skewness $S_k = m_3/m_2^{3/2}$ is given by (5). ■

References

Abdulai, A., 2001, Technical Efficiency during Economic Reform in Nicaragua: Evidence from Farm Household Survey Data, Economic Systems 25,113-125

Apezteguia, B. I. and Garate, M. R., 1997, Technical efficiency in the Spanish agrofood Industry, Agricultural Economics, 17, 179-189.

Audibert, M., 1997, Technical Inefficiency Effects Among Paddy Farmers in the Villages of the 'Office du Niger' Mali, West Africa, Journal of Productivity Analysis, 8, 379-394.

Battese, G. E. and Coelli, T. J., 1988, Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data, Journal of Econometrics, 38, 387 - 399.

Battese, G. E. and Coelli, T. J., 1993, A Stochastic Frontier Production Function Incorporating a Model for Technical Inefficiency Effects, Working Papers, No. 69, in Econometrics and Applied Statistics, Department of Economics, University of New England.

Battese, G. J. and Coelli, T. J., 1995, A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data, Empirical Economics 20, 325 - 332.

Bhattacharyya, AN., Bhattacharyya, AR., and Kumbhakar, S. C.,1996, Government Interventions, Market Imperfections, and Technical Inefficiency in a Mixed Economy: A Case Study of Indian Agriculture, Journal of Comparative Economics, 22, 219-241.

Cullinane, K., Song, D-W. and Gray, R., 2002, A Stochastic Frontier Model of the Efficiency of Major Container Terminals in Asia: Assessing the Influence of Administrative and Ownership Structures, Transportation Research Part A, 36, 743-762.

Delorme,C.D. , Thompson,H.G. and Warren,R.,1995, Money and Production: A Stochastic Frontier Approach, Journal of Productivity Analysis, 6, 333-342.

Ferrantino,M.J. and Ferrier,G.D.,1995, The Technical Efficiency of Vacuum-pan Sugar Industry of India: An Application of a Stochastic Frontier Production

Function Using Panel Data, European Journal of Operational Research,80,639-653

Hattori, T.,2002, Relative Performance of U.S. and Japanese Electricity Distribution: An Application of Stochastic Frontier Analysis, Journal of Productivity Analysis, 18, 269-284.

Kaynak, H. and Pagan, J. A. , 2003, Just-in-Time Purchasing and Technical Efficiency in the US Manufacturing Sector, International Journal of Production Research, 41, 1-14.

Mercedes G-A., 2000, Efficiency and Technical Progress: Sources of Convergence in the Spanish Regions, Applied Economics, 32,467-478.

Murillo-Zamorano,L.R. and Vega-Cervera, J.A., The Use of Parametric and Non-Parameteric Frontier Methods to Measure the Productive Efficiency in the Industrial Sector: A Comparative Study, International Journal of Production Economics, 69, 265-275.

Morrison,C., Morrison,P. and Frengley; G.A.G.,2000,Efficiency in New Zealand sheep and Beef Farming:The Impacts of Regulatory Reform, The Review of Economics and Statistics,82,325-337.

Pascoe, S. and Coglan, L., 2002, The Contribution of Unmeasurable Inputs to Fisheries Production: An Analysis of the Technical Efficiency of Fishing Vessels in the English Channel, American Journal of Agricultural Economics, 84, 585-597.

Puig-Junoy, J., 2001, Technical Inefficiency and Public Capital in U.S. States: A Stochastic Frontier Approach, Journal of Regional Science, 41, 75-96.

Seyoum,E.T. , Battese, G.E. and Fleming, E.M., 1998,Technical Efficiency and Productivity of Maize Producers in Eastern Ethiopia: A Study of Farmers Within and Outside the Saskawa-Global 2000 project, Agricultural Economics, 19,341-348.

Shao,B.B.M, and Lin, W.T., 2001, Measuring the Value of Information Technology in Technical Efficiency with Stochastic Production Frontier, Information and Software Technology 43,447-456.

Sharma, H. R. , Leung,P.S. and Zaleski,1 H. M. , 1997, Productive Efficiency of the Swine Industry in Hawaii: Stochastic Frontier vs. Data Envelopment Analysis, Journal of Productivity Analysis, 8, 447-459.

Sharma,K.R. and Leung, P.S., 2000,Technical Efficiency of Carp Production in India: A Stochastic Frontier Production Function Analysis, Aquaculture Research,31,937-947.

Tian,W., 2000, Technical Inefficiency and Its Determinants in China's Grain , Production, Journal of Productivity Analysis, 13, 159-174.

Tsukuda, Y. and Miyakoshi, T., 2002, A Re-Consideration on Technical Inefficiency in Stochastic Frontier Production Models, Institute of Policy and Planning Sciences, Tsukuba University, Discussion Paper Series, No.993

Tsukuda, Y. and Miyakoshi, T., 2003, An Alternative Method for Predicting Technical Inefficiency in Stochastic Frontier Models, Applied Economics Letters, Forthcoming.

Uri, N. D., 2001, The Effect of Incentive Regulation on Productive Efficiency in Telecommunications, Journal of Policy Modeling, 23, 825-846.

Wadud, A. and White, B., 2000,Farm Household Efficiency in Bangladesh: A Comparison of Stochastic Frontier and DEA methods, Applied Economics, 32,1665-1673.

Wang, W., 2000, Evaluating the Technical Efficiency of Large US Law Firms, Applied Economics, 32, 689-695.

Wilson,P., Hadley,D. and Asby, C. 2001,The Influence of Management Characteristics on the Technical Efficiency of Wheat Farmers in Eastern England, Agricultural Economics, 24, 329-338.

Yao,S.,Liu,Z. and Zhang,Z.,2001, Spatial Differences of Grain Production Efficiency in China,1987-1992, Economics of Planning, 34,139-157.

Table1. Numerical Comparison of Predictions

μ^*	$\sigma^* = 1.0$			$\sigma^* = 0.6$			$\sigma^* = 0.3$			$\sigma^* = 0.2$		
	ETIE	MTIE	SKEW	ETIE	MTIE	SKEW	ETIE	MTIE	SKEW	ETIE	MTIE	SKEW
0.0	0.477	0.491	−0.112	0.343	0.333	0.232	0.201	0.183	0.561	0.142	0.126	0.691
0.25	0.514	0.540	−0.251	0.398	0.401	0.021	0.282	0.279	0.167	0.242	0.242	0.113
0.5	0.554	0.592	−0.406	0.461	0.480	−0.220	0.391	0.404	−0.276	0.384	0.394	−0.415
1.0	0.640	0.699	−0.770	0.604	0.645	−0.792	0.615	0.632	−0.883	0.625	0.632	−0.614
1.5	0.727	0.795	−1.227	0.739	0.778	−1.434	0.767	0.777	−0.949	0.772	0.777	−0.614
2.0	0.808	0.868	−1.805	0.838	0.865	−1.961	0.858	0.865	−0.950	0.862	0.865	−0.614

Note: The enties are calculated by EXCEL.
ETIE, MTIE and SKEW denote the conditional expectation and median, and skewness, respectively.

Table 2. The estimates of σ* in the previous researches

Articles	$\sigma^2 = \sigma_V^2 + \sigma_U^2$	$\gamma = \sigma_U^2 / \sigma^2$	$\sigma^* \equiv \sqrt{\gamma(1-\gamma)\sigma^2}$
1. Uri(2001,pp.841,Table 3): Truncated normal	13.977	0.972	0.621
2. Kaynak et al (2003, pp.8, Table1):Truncated normal	0.880	0.283	0.423
3. Pascoe et al (2002, pp. 592, Table 3):Truncated normal	0.844	0.762	0.391
4. Cullinance,et al(2002,pp. 756, Table 4):Truncated normal	0.457	0.481	0.338
5. Sharma, et al(1997,pp. 452): Truncated normal	0.900	0.870	0.320
6. Audibert(1997,pp.387, Table 2,Model3):Truncated normal	0.389	0.852	0.221
7. Battese&Coelli(1995,pp. 329): Truncated normal	0.740	0.952	0.184
8. Bhattacharyya,et al (1996, pp.237,Table 3):Half-normal	0.091	0.727	0.135
9. Murillo-Zamorano, et al (2001 , pp.270, Table 2): Truncated normal	0.120	0.829	0.130
10. Wang(2000,pp.694,Table3): Half-normal	0.076	0.660	0.130

Notes: Some papers cited above provide several models of the technical inefficiency distributions.

Table 3. Difference between the ETIE and MTIE

Articles	ETIE (1)	μ* (2)	SKEW (3)	MTIE (4)	Relative difference ((1.B)-(4.B))/(1.B) (5)	((1.B)-(4.B))/((1.B)-(1.A)) (6)
Kaynak et al (2003,pp.8):						
Min(A)	0.019	___	___	___		
Median(B)	0.082	-1.825	1.353	0.062	0.24	0.32
Max(C)	0.465	0.632	-0.390	0.487		
Pascoe et al (2002,pp. Fig.1):						
Min(A)	0.050	___	___	___		
Mode(B)	0.150	-0.574	0.963	0.124	0.17	0.26
Max(D)	0.651	1.125	-1.054	0.676		
Uri(2001):						
Min(A)	0.098	-3.298	1.208	0.075		
Mean(B)	0.251	-0.638	0.621	0.222	0.12	0.19
Max(C)	0.703	1.362	-1.244	0.747		
Cullinance,et al (2002, pp.757,Table5):						
Min(A)	0.195	-0.135	0.678	0.168		
Median(B)	0.278	0.193	0.248	0.271	0.03	0.08
Max(C)	0.377	0.453	-0.163	0.388		
Sharma, et al (1997,pp.453,Table4):						
Min(A)	0.098	-0.750	1.214	0.076		
Mean(B)	0.251	0.136	0.343	0.240	0.04	0.07
Max(C)	0.703	1.266	-1.005	0.718		

Notes:

a) Uri (2001) does not report any technical inefficiency. We use the values of Sharma, et al (1997) as proxies because both studies deal with the USA industry.

b) The line "___" in the entry means that EXCEL does not produce the proper values because of the extreme quantile value of the standard normal distribution.

c) The (1.B) means the column (1) and the row (B).