

A Monte Carlo Analysis for Stochastic Distance Function Frontier

Tao Zhang

Macao Polytechnic Institute, School of Public Administration
Rua Luis Gonzaga Gomes, Macao, China
e-mail. taozhang7608@hotmail.com)

crossref <http://dx.doi.org/10.5755/j01.ee.23.3.1937>

Being different from a Data Envelopment Analysis (DEA) whose results can be decomposed, the parametric stochastic frontier method can not deal with multiple outputs and multiple inputs easily and directly. Although stochastic cost frontier and stochastic revenue frontier can potentially be used to describe a multi-output, multi-input production, these functions only allow the possibility of specifying a multi-product technology when price information is available and some required behavioral assumptions are satisfied. An alternative for cost or revenue function called distance functions was developed by Shephard (1953). Fare et al. (1993) first used parametric linear programming (PLP) approach to estimate a translog distance function. In 1996, Tim Coelli and Sergio Perelman (1996) applied this approach (called stochastic distance function frontier) to adjust the function form of stochastic frontier and make it suitable for multi-product analysis. Since then, the stochastic distance function frontier has begun to be widely used in particular applications. But, in all these studies, the stochastic distance function frontier approach required changing the function form and thus it made the influence of function assumption on results stronger. Some econometricians proposed that the distance function approach in handling multi-product stochastic frontier might introduce regressor endogeneity and induce estimator inconsistency in estimation. It is also proposed that Multi-output stochastic distance functions suffer from input-output separability and linear homogeneity in outputs. However, we can not observe any studies on Monte Carlo test of stochastic distance function frontier. The principle purpose of this study is to contribute to the methodology discussion of stochastic distance function frontier using Monte Carlo analysis for multi-product efficiency measurement. Thus, a Monte Carlo analysis for stochastic distance function frontier model was developed to test the asymptotic properties. The technology for the framework of stochastic distance function, used to overcome the criticisms related to the stochastic frontier approach, is specified as a translog function. Then, the basic method to estimate the stochastic frontier is provided and the maximum likelihood function is also given. In the Monte Carlo experiment, 1000 replications are set for analysis. The results show that, except for the scenario with equal outputs, stochastic distance function frontier will yield biased estimators even with large sample size. The 2-output model will give better estimators than 3-output model. An increasing sample size will improve the relative performance of ML estimations for stochastic distance function frontier. Therefore, the Monte Carlo analysis indicates the result that the stochastic distance function frontier is probably biased for multi-output production.

Keywords: *Stochastic distance function frontier; Monte Carlo analysis; Bias; Estimator.*

Introduction

Farrell (1957) and Debreu (1951) first created the basic efficiency index models. These models can be regarded as applications of Shephard's (1953, 1970) and Malmquist's (1953) distance functions. Much of the previous work for 'traditional' DEA (Data Envelopment Analysis) normally used input and output distance functions and radial contractions or expansions of inputs and outputs. An appealing property of the DEA approach is that multiple-input and multiple-output technologies can be modelled without behavioral assumptions and price information. Nevertheless, they have some limitations which arise from their 'deterministic nature'. (Xue & Harker, 1999), (Hirschberg & Lloyd, 2002). Being different from the Data Envelopment Analysis (DEA) whose results can be decomposed (Zhang, 2010a), the parametric stochastic frontier method can not deal with multiple outputs and multiple inputs easily and directly (Zhang et al., 2008), (Zhang, 2009, 2010b). Although stochastic cost frontier and stochastic revenue frontier can potentially be used to describe a multi-output, multi-input production, these

functions only allow the possibility of specifying a multi-product technology when price information is available and some required behavioral assumptions are satisfied.

An alternative for cost or revenue function called distance functions was developed by Shephard (1953). Fare et al. (1993) first used parametric linear programming (PLP) approach to estimate a translog distance function. In 1996, Tim Coelli and Sergio Perelman (1996) applied this approach (called stochastic distance function frontier) to adjust the function form of stochastic frontier and make it suitable for multi-product analysis. Since then, the stochastic distance function frontier has begun to be widely used in particular applications. Tim Coelli and Sergio Perelman (2000) applied stochastic distance function frontier in analyzing multi-outputs technical efficiency of European railways. Xavier Irz and Colin Thirtle (2004) investigated the technological development in agriculture using stochastic input distance function frontier. Carol Newman and Alan Matthews (2006) measured the productivity performance of Irish dairy farms using stochastic output distance function frontier.

However, in all these studies, the stochastic distance function frontier approach required changing the function form and thus it made the influence of function assumption on results stronger. Some econometricians proposed that the distance function approach in handling multi-product stochastic frontier might introduce regressor endogeneity and induce estimator inconsistency in estimation (Grosskopf et al., 1997). It is also proposed that Multi-output stochastic distance functions suffer from input-output separability and linear homogeneity in outputs. The principle aim of this study is to contribute to the stock of knowledge of the estimation method for multi-product efficiency measurement.

Currently, the only observed paper which tested the methodology issue of stochastic distance function frontier was written by Tim Coelli and Sergio Perelman (1999). In this paper, they investigated the sensitivity of estimated technical efficiency scores from different methods including stochastic distance function frontier. However, it only compared the technical efficiency scores from multi-output production function estimated by the parametric linear programming approach, multi-output orientated DEA, and C-OLS distance function frontier. This paper did not investigate the asymptotic characters of stochastic distance function frontier approach. Thus, hitherto, we can not observe any studies on Monte Carlo analysis of stochastic distance function frontier. This paper hopes to shed some new light on the methodology discussion of stochastic distance function frontier using Monte Carlo analysis in multi-output efficiency measurement.

The Tanslog Distance Functions

Fare et al. (1985, 1993, and 1994) introduced the concept of using distance functions to express the output bundle of a multiple product technology. Then, the framework of stochastic distance function has been developed in the literature to overcome the criticisms related to multiple product technology of the stochastic frontier approach. The technology is specified as a translog function (Zhang & Xue, 2005). Coelli and Perelman (1996) provided discussions of econometric estimation of stochastic distance functions with multiple outputs.

The value of the distance function can not be observed and it can not be estimated directly either. Lovell *et al.* (1994) suggested a convenient approach circumventing this problem using the property of linear homogeneity of the distance function. It should be noted that homogeneity implies:

$$Di(x, \beta y) = \beta Di(x, y), \text{ for any } \beta > 0 \tag{1}$$

Here, *y* is a vector of outputs of dimension L and setting $\beta=1/Y_L$, where arbitrarily chosen Y_L denotes its first component, the previous equation can be expressed in logarithmic form as:

$$\text{Ln}Di(x, y) = \text{Ln}Y_L \text{Ln}Di(x, y/Y_L) \tag{2}$$

Then, above equation can be rewritten as:

$$\text{Ln}Di(x, y) - \text{Ln}Y_L = \text{Ln}Di(x, y/Y_L) \tag{3}$$

In addition, it is assumed that:

$$\text{Ln}Di(x, y) = -U_i \tag{4}$$

Combining Equations (3) and (4) to give:

$$-\text{Ln}Y_L = \text{Ln}Di(x, y/Y_L) + U_i \tag{5}$$

In the empirical application, the translog production function form was chosen. The translog production function form of $\text{Ln}Di(x, y)$ can be written as:

$$\begin{aligned} \text{Ln}D(x, y) = & \alpha + \sum_{k=1}^K \beta_k \text{Ln}X_k + \sum_{l=1}^L \alpha_l \text{Ln}Y_l + \frac{1}{2} \sum_{k=1}^K \sum_{k'=1}^K \beta_{kk'} \text{Ln}X_k \text{Ln}X_{k'} + \\ & + \frac{1}{2} \sum_{l=1}^L \sum_{l'=1}^L \alpha_{ll'} \text{Ln}Y_l \text{Ln}Y_{l'} + \sum_{k=1}^K \sum_{l=1}^L \xi_{kl} \text{Ln}X_k \text{Ln}Y_l \end{aligned} \tag{6}$$

where *i* refers to the *i*th unit, Ln represents the natural logarithm, *Y* are the output variables and *X* are the input variables. α, β, ξ are parameters to be estimated.

Here, the homogeneity restrictions are:

$$\begin{aligned} \sum_{l=1}^L \alpha_l &= 1 \\ \sum_{l=1}^L \alpha_{ll'} &= 0 \\ \sum_{l=1}^L \xi_{kl} &= 0 \end{aligned}$$

and required conditions for symmetry are:

$$\begin{aligned} \beta_{kk'} &= \beta_{k'k} \\ \alpha_{ll'} &= \alpha_{l'l} \end{aligned}$$

Then according to equation (1), we can give:

$$\begin{aligned} \text{Ln}(D(x, y) / Y_L) = & \alpha + \sum_{k=1}^K \beta_k \text{Ln}X_k + \sum_{l=1}^{L-1} \alpha_l \text{Ln}(Y_l / Y_L) + \\ & + \frac{1}{2} \sum_{k=1}^K \sum_{k'=1}^K \beta_{kk'} \text{Ln}X_k \text{Ln}X_{k'} + \frac{1}{2} \sum_{l=1}^{L-1} \sum_{l'=1}^{L-1} \alpha_{ll'} \text{Ln}(Y_l / Y_L) \text{Ln}(Y_{l'} / Y_L) + \\ & + \sum_{k=1}^K \sum_{l=1}^{L-1} \xi_{kl} \text{Ln}X_k \text{Ln}(Y_l / Y_L) \end{aligned} \tag{7}$$

Finally, the function form can be rewritten as

$$\begin{aligned} -\text{Ln}(Y_L) = & \alpha + \sum_{k=1}^K \beta_k \text{Ln}X_k + \sum_{l=1}^{L-1} \alpha_l \text{Ln}(Y_l / Y_L) + \\ & + \frac{1}{2} \sum_{k=1}^K \sum_{k'=1}^K \beta_{kk'} \text{Ln}X_k \text{Ln}X_{k'} + \frac{1}{2} \sum_{l=1}^{L-1} \sum_{l'=1}^{L-1} \alpha_{ll'} \text{Ln}(Y_l / Y_L) \text{Ln}(Y_{l'} / Y_L) + \\ & + \sum_{k=1}^K \sum_{l=1}^{L-1} \xi_{kl} \text{Ln}X_k \text{Ln}(Y_l / Y_L) + U \end{aligned} \tag{8}$$

Again, the technical efficiency can be calculated as: $\text{TE} = \text{EXP}(-U) = \text{Di}(x, y)$, where TE changes from 0 to 1, and $\text{Ln}Di(x, y) \leq 0$.

If all the outputs are equals in the value, then function (8) can be simplified to the common one-output production function as:

$$\begin{aligned} -\text{Ln}(Y_L) = & \alpha + \sum_{k=1}^K \beta_k \text{Ln}X_k + \\ & + \frac{1}{2} \sum_{k=1}^K \sum_{k'=1}^K \beta_{kk'} \text{Ln}X_k \text{Ln}X_{k'} + U \end{aligned} \tag{9}$$

The stochastic frontier analysis

The stochastic production function frontier model was first proposed by Aigner, Lovell and Schmidt (1977). As this basic model has been widely used and our purposes

are focused on distance functions instead of specification assumptions for regressions, this analysis is confined to the basic half-normal specification of inefficiency components.

The model can be defined as:

$$y_i = \beta x_i + V_i - U_i, \quad i = 1, 2, \dots, N, \quad (10)$$

Where y is output; x is a vector of inputs; β is a vector of parameters. The stochastic frontier is specified by decomposing the error term into two components respectively accounting for technical inefficiency (U_i) and random error (V_i). The basic stochastic frontier assumes that the U_i are half-normal and independent of the V_i . The random errors are assumed to be normal. The method of maximum likelihood is used to estimate the unknown parameters. Battese and Corra (1977) propose the loglikelihood function as:

$$\log(L) = -\frac{N}{2} \log(\pi/2) - \frac{N}{2} \log(\sigma^2) + \sum_{i=1}^N \log\{1 - \Phi[\frac{(y_i - x_i \beta)}{\sigma} \sqrt{\frac{\gamma}{1-\gamma}}]\} - \frac{1}{2\sigma^2} \sum_{i=1}^N (y_i - x_i \beta)^2 \quad (11)$$

where $\Phi(\cdot)$ is distribution function of a standard normal random variable, $\sigma^2 = \sigma_U^2 + \sigma_V^2$ and $\gamma = \sigma_U^2 / \sigma^2$. The γ , β and σ are estimated by finding the maximum value of likelihood function defined in above equation. In addition, the mean technical efficiency can be estimated by:

$$\text{mean TE} = 2 \cdot \exp(\gamma \sigma^2 / 2) [1 - \Phi(\sqrt{\gamma} \sigma)] \quad (12)$$

It should be noted that this estimated mean technical efficiency is only applicable when y is the logarithm of production (Battese and Coelli, 1988).

Monte Carlo Analysis Design

In this experiment, 1000 replications are set for our Monte Carlo analysis. The sample space includes α , β , ξ , γ , σ , N and X , where N is sample size and X is a vector of inputs. We make sample size N changed from 50 to 800. From the results of Tim Coelli (1995), it is showed that the γ bias diminishes as the value of γ approaches one. Because our study only concerns distance functions instead of stochastic frontier regressions, we fix γ at 0.96 and σ at 0.5. We reduce the number of 'X's to only two variables, one is constant and the other is the input generated from uniform (5,15). We generate outputs from function(6) instead of common function (10), because the final stochastic frontier distance function(8) is transformed from translog distance function(6) and therefore common stochastic frontier function(10) can not provide multiple outputs in our experiments. To test multi-output distance function, we will make both 2-output experiments and 3-output experiments. The true parameters for both 2-output and 3-output experiments are listed as follows:

Table 1

The true values of parameters

	The true parameters for 2-output experiments	The true parameters for 3-output experiments
α_1	0.5	1/3
α_2	0.5	1/3
α_3		1/3
α_{11}	0.2	0.2

	The true parameters for 2-output experiments	The true parameters for 3-output experiments
α_{12}	-0.2	-0.1
α_{13}		-0.1
α_{22}	0.2	0.2
α_{23}		-0.1
α_{33}		0.2
α	1	1
β_1	0.5	0.5
β_{11}	0.25	0.25
ξ_{11}	0.1	0.1
ξ_{12}	-0.1	-0.05
ξ_{13}		-0.05

In function(6), we also have to provide the technical inefficiency components $\text{LnDi}(x, y)$ and random error components V_i . In each of the replications, two vectors of standard normal random numbers with length 800 are generated to construct the technical inefficiency components and random error components. Here, the V_i are generated by multiplying the first vector of random numbers by the σ_V and the U_i are formed by multiplying the absolute values of the second vector of random numbers by the σ_U . It should be noted that σ_V and σ_U were obtained from given γ and σ . Through functions $\sigma^2 = \sigma_U^2 + \sigma_V^2$ and $\gamma = \sigma_U^2 / \sigma^2$, we can find that σ_V is fixed at 0.1 and σ_U is $\sqrt{0.24}$. Based on the function (6), we assume that all the outputs are positive. This assumption gives additional constraints on the output values generated from function (6).

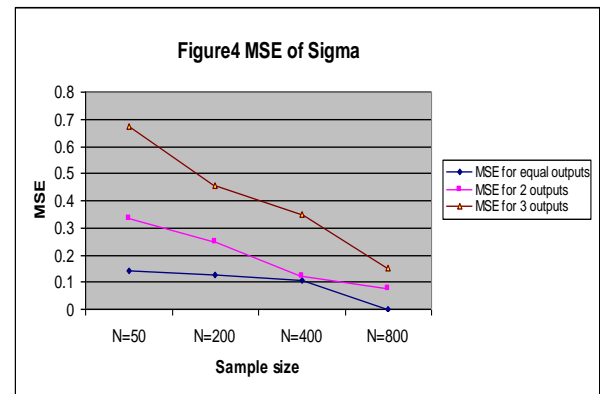
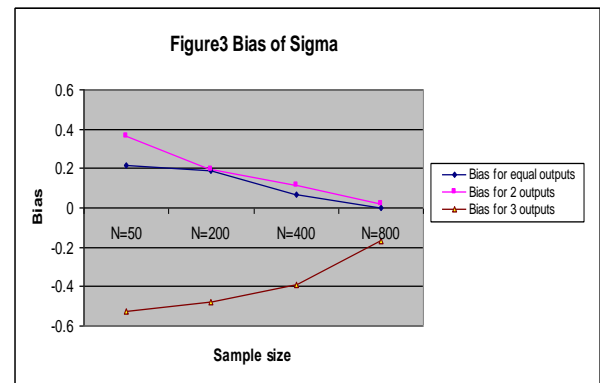
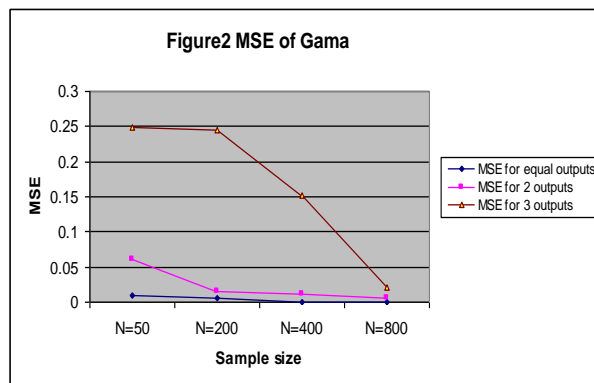
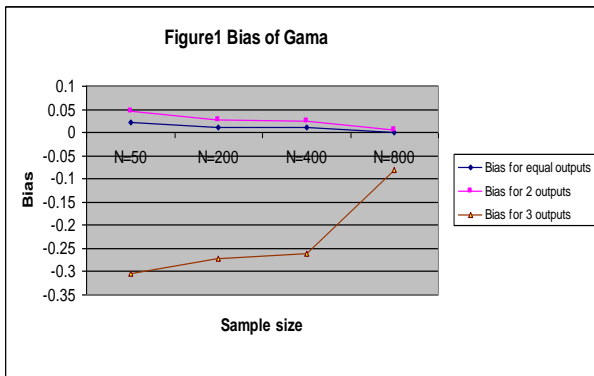
Furthermore, we can make different experiments for different combinations of the outputs. Thus, we will have three scenarios individually for equal outputs, 2 outputs and 3 outputs.

Results of the Monte Carlo Experiments

From function (12), it is obvious that the mean technical efficiency (TE) will be decided by both γ and σ . Coelli (1995) proposed that the bias in the estimator of γ will provide a misleading indication of the level of mean TE. The positive bias in the γ estimate when its true value is zero results in negative bias in the estimate of TE. And, the bias in σ will influence the size of the bias in the TE estimates. Thus, if both γ and σ are unbiased, then the mean technical efficiency estimate is unbiased too. Therefore, in this study, we only report briefly the bias and MSE (Mean Squared Error) of γ and σ estimators from the results of Monte Carlo experiments. The bias and mean squared error of γ and σ estimators are depicted in figure1, 2, 3 and 4. The absolute values of the biases and MSEs of both γ and σ for the equal-output scenario are the smallest in three scenarios. It is also showed that the bias and MSE of γ for the equal-output scenario are always lower than those of σ . However, they all approach 0 when the sample size is 800. This result corresponds with the find discovered by Coelli (1995) who proposes that the bias diminishes to zero when γ approaches one and sample size increases to 800. (The estimator biases for the equal-output

scenario in our study are positive, but the biases for the one-output experiments in Coelli's (1995) study are negative when γ approaches one. The reason is that the equal-output function (8) is deduced from the distance function. And, it has negative sign on the dependent variable $\ln(Y_L)$.

The absolute values of the biases and MSEs of γ and σ for the 2-output scenario are all higher than those of the equal-output scenario and lower than those of the 3-output scenario. The estimator biases for the 3-output scenario are negative and the absolute values of them are the largest in three scenarios. All above results indicate that the stochastic distance function frontier model with fewer outputs probably has less biased estimators. More outputs means more bias in estimators and estimated mean technical efficiency.



Conclusions

Except for the scenario with equal outputs, stochastic distance function frontier will yield biased estimators even with large sample size. The 2-output model will give better estimators than 3-output model. An increasing sample size will improve the relative performance of ML estimations for stochastic distance function frontier. Therefore, the Monte Carlo analysis finally shows the result that the stochastic distance function frontier is probably biased for multi-output production.

References

- Aigner, D., Lovell, K., & Schmidt, P. (1977). Formulation and Estimation of Stochastic Frontier Models. *Journal of Econometrics*, 6(1), 21-37. [http://dx.doi.org/10.1016/0304-4076\(77\)90052-5](http://dx.doi.org/10.1016/0304-4076(77)90052-5)
- Battese, G. E., & 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. [http://dx.doi.org/10.1016/0304-4076\(88\)90053-X](http://dx.doi.org/10.1016/0304-4076(88)90053-X)
- Battese, G. E., & Corra, G. S. (1977). Estimation of Production Frontier Model: with Application to the Pastoral Zone of Eastern Australia. *Australian Journal of Agricultural Economics*, 21, 169-179. <http://dx.doi.org/10.1111/j.1467-8489.1977.tb00204.x>
- Coelli, T. (1995). Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis. *Journal of Productivity Analysis*, 6, 247-268 <http://dx.doi.org/10.1007/BF01076978>
- Coelli, T. J., & Perelman, S. (1996). Efficiency Measurement, Multiple-Output Technologies and Distance Functions: With Application to European Railways, CREPP Working Paper 96/05. University of Liege.
- Coelli, T. J., & Perelman, S., (2000). Technical Efficiency of European Railways: A Distance Function Approach. *Applied Economics*, 32(15), 1967-1976. <http://dx.doi.org/10.1080/00036840050155896>

- Coelli, T., & Perelman, S. (1999). A Comparison of Parametric and Non-Parametric Distance Functions: With Application to European Railways. *European Journal of Operational Research*, 117(2), 326-339. [http://dx.doi.org/10.1016/S0377-2217\(98\)00271-9](http://dx.doi.org/10.1016/S0377-2217(98)00271-9)
- Debreu, G. (1951). The Coefficient of Resource Utilization. *Econometrica*, 19(3), 273-292. <http://dx.doi.org/10.2307/1906814>
- Fare, R., Grosskopf, S., & C. A. K. Lovell. (1985). *The Measurement of Efficiency of Production*, Boston, Kluwer.
- Fare, R., Grosskopf, S., & C. A. K. Lovell. (1994). *Production Frontiers*. Cambridge University Press, U.K.
- Fare, R., Grosskopf, S., Lovell, C. A. K., & Yaisawarng, S. (1993). Derivation of Shadow Prices for Undesirable Outputs: A Distance Function Approach. *Review of Economics and Statistics*, 75(2), 374-380. <http://dx.doi.org/10.2307/2109448>
- Farrell, M. J. (1957). The Measurement of Production Efficiency. *Journal of the Royal Statistical Society Series A* 120, 253-290 <http://dx.doi.org/10.2307/2343100>
- Grosskopf, S., Hayes, K., Taylor, L., & Weber, W. (1997). Budget Constrained Frontier Measurement of Fiscal Equality and Efficiency in Schooling. *Review of Economics and Statistics*, 79(1), 116-124. <http://dx.doi.org/10.1162/003465397556458>
- Hirschberg, J. G., Lloyd, P. J., (2002). Does the Technology of Foreign-Invested Enterprises Spill Over to Other Enterprises in China? An Application of Post-DEA Bootstrap Regression Analysis. In: Lloyd, P.J., Zang, X.G. (Eds.), *Modeling the Chinese Economy*. Edward Elgar Press; London
- Irz, X., & Thirtle, C. (2004). Dual Technological Development in Botswana Agriculture: A Stochastic Input Distance Function Approach. *Journal of Agricultural Economics*, 55(3), 455-478 <http://dx.doi.org/10.1111/j.1477-9552.2004.tb00110.x>
- Lovell, C. A. K., Richardson, S., Travers, P., & Wood, L.L. (1994). Resources and Functionings: A New View of Inequality in Australia, W. Eichhorn (ed.), *Models and Measurement of Welfare and Inequality*, Berlin, Springer-Verlag, 787-807
- Malmquist, S. (1953). Index Numbers and Indifference Surfaces. *Trabajos de Estadística*, 4, 209-242. <http://dx.doi.org/10.1007/BF03006863>
- Newman, C., & Matthews, A. (2006). The Productivity Performance of Irish Dairy Farms 1984–2000: a Multiple Output Distance Function Approach. *Journal of Productivity Analysis*, 26, 191-205 <http://dx.doi.org/10.1007/s1123-006-0013-7>
- Shephard, R. W. (1953). *Cost and Production Functions*, Princeton University Press, Princeton, NJ.
- Shephard, R. W. (1970). *Theory of Cost and Production Functions*, Princeton: Princeton University Press.
- Xue, M., & Harker, P. T. (1999). Overcoming the Inherent Dependency of DEA Efficiency Scores: A Bootstrap Approach. Wharton Financial Institutions Centre, Working Paper, 99-17.
- Zhang, T. (2009). Frame Work of Data Envelopment Analysis-A Model to Evaluate the Environmental Efficiency of China's Industrial Sectors. *Biomedical and Environmental Sciences*, 21, 8-13 [http://dx.doi.org/10.1016/S0895-3988\(09\)60015-5](http://dx.doi.org/10.1016/S0895-3988(09)60015-5)
- Zhang, T. (2010a). Decomposing Allocative Efficiency for Multi-Production Systems, *E+M Ekonomie a Management-Economics and Management*, XIII(2), 71-77
- Zhang, T. (2010b). Environmental Performance Assessment of China's Manufacturing. *Asian Economic Journal*, 24(1), 45-68. <http://dx.doi.org/10.1111/j.1467-8381.2010.02026.x>
- Zhang, T., & Eoghan, G. (2008), A Comparative Analysis of Multi-Output Frontier Models. *Journal of Zhejiang University - Science A* 9(10), 1426-1436
- Zhang, T., & Xue B. (2005). Environmental Efficiency Analysis of China's Vegetable Production. *Biomedical and Environmental Sciences*, 18(1), 21-30

Tao Zhang

Monte Karlo analizė Stochastinei ribinei atstumo funkcijai

Santrumpa

Farrell (1957) ir Debreu (1951) pirmieji sukūrė pagrindinių našumo indeksų modelius. Šiuos modelius galima laikyti Duomenų paketo analizės (toliau DEA) papildymais. (*plg. angl. Data Envelopment Analysis*). Patraukli DEA metodo ypatybė yra ta, kad daugkartinių sąnaudų ir daugkartinės išėigos technologijos gali būti modeliuojamos be prielaidų apie elgesį ir informacijos apie kainą. Tačiau jie turi kai kuriuos apribojimus, atsirandančius iš jų „deterministinės prigimties“. Kadangi skiriasi nuo Duomenų paketo analizės, parametrinis stochastinis ribinis metodas negali sietis su daugkartinėmis sąnaudomis ir daugkartinė išėiga. Nors stochastinė kainos riba ir stochastinė pajamų riba gali būti potencialiai naudojamos nusakyti daugkartinės išėigos ir daugkartinių sąnaudų našumą, šios funkcijos tik suteikia galimybę nustatyti daugkartinio gaminio technologiją, kai galima gauti informaciją apie kainą.

Alternatyva kainos ir pajamų funkcijai yra *atstumo* funkcija. Ją išplėtojo Shephard (1953). 1996 metais, Tim Coelli ir Sergio Perelman (1996) pritaikė šį metodą (vadinamą *stochastine ribine atstumo funkcija*), kad suderintų stochastinės ribos funkcijos formą ir padarytų ją tinkamą analizuojant daugkartinį gaminį. Nuo tada stochastinė ribinė atstumo funkcija buvo pradėta plačiai naudoti specifiniams pritaikymams. Tačiau visuose darbuose stochastinės ribinės atstumo funkcijos metodas reikalavo pakeisti funkcijos formą ir tokiu būdu sustiprino funkcijos prielaidų įtaką rezultatams. Kai kurie ekonometrikos specialistai teigė, kad atstumo funkcijos metodas valdant daugkartinio gaminio stochastinę ribą galėtų įvesti regresoriaus endogeniškumą

ir sukelti nesuderinamumą įvertinime (Grosskopf ir kt., 1997). Taip pat teigiama, kad daugkartinės išiegos stochastinės atstumo funkcijos nukenčia dėl sąnaudų-šieigos atskiriamumo ir linijinio vientisumo išiegoje. Pagrindinis šio tyrimo tikslas yra papildyti turimus duomenis apie kaupimo įvertinimo metodą daugkartinio gaminio našumui matuoti.

Darbu rašytų minėta tema yra vienas kitas. Stochastinės ribinės atstumo funkcijos metodologijos problemą tyrė Tim Coelli ir Sergio Perelman (1999). Šiame darbe jie tyrinėjo vertinamo techninio našumo rezultatų jautrumą skirtingais metodais, įskaitant stochastinės ribinės atstumo funkcijas. Tačiau jie tik palygino daugkartinės išiegos našumo funkcijas, įvertintas parametrinio linijinio programavimo metodu, taip pat daugkartinę išiega nukreiptą į DEA, ir C-OLS atstumo ribinę funkciją. Šiame darbe nenagrinėtos scholastinės ribinės atstumo funkcijos metodo asimptotinės savybės. Tokiu būdu, iki šiol mes negalėjome stebėti jokios stochastinės ribinės atstumo funkcijos Monte Karlo analizės tyrimų. Šiuo darbu tikimasi atskleisti stochastinės ribinės atstumo funkcijos metodologiją, naudojant Monte Karlo analizę daugkartinės išiegos našumo matavimuose.

Fare ir kt. (1993) pristatė atstumo funkcijų panaudojimo koncepciją norėdami atskleisti daugkartinio gaminio technologijos išiegos paketą. Vėliau mokslinėje literatūroje buvo išplėtotas stochastinės atstumo funkcijos struktūra, siekiant įveikti kritika, susijusią su stochastinės ribos metodo daugkartinio gaminio technologija. Technologija yra apibrėžta kaip *translog funkcija*. Coelli ir Perelman (1996) paruošė diskusijas apie stochastinių atstumo funkcijų su daugkartine išiega ekonometrinius įvertinimus. Atstumo funkcijos vertė gali būti nestebima ir neįvertinama tiesiogiai. Lovell ir kt. (1994) pasiūlė patogų metodą, kaip apeiti šią problemą naudojant atstumo funkcijos linijinio vientisumo savybę. Taikant empiriškai, buvo pasirinkta *translog* našumo funkcijos forma. *Translog* našumo funkcijos formą $LnDi(x, y)$ galima užrašyti taip

$$LnD(x, y) = \alpha + \sum_{k=1}^K \beta_k LnX_k + \sum_{l=1}^L \alpha_l LnY_l + \frac{1}{2} \sum_{k=1}^K \sum_{k'=1}^K \beta_{kk'} LnX_k LnX_{k'} + \frac{1}{2} \sum_{l=1}^L \sum_{l'=1}^L \alpha_{ll'} LnY_l LnY_{l'} + \sum_{k=1}^K \sum_{l=1}^L \xi_{kl} LnX_k LnY_l \quad (6)$$

Kur i žymi i -tąjį vieneta, Ln žymi natūralų logaritmą, Y yra išiegos kintamieji ir X yra sąnaudų kintamieji. α, β, ξ yra parametrai, kuriuos reikia įvertinti.

Atlikus dedukcinį patikrinimą, mes pagaliau galime gauti funkcijos formą, kurią galima užrašyti taip

$$- Ln(Y_L) = \alpha + \sum_{k=1}^K \beta_k LnX_k + \sum_{l=1}^{L-1} \alpha_l Ln(Y_l / Y_L) + \frac{1}{2} \sum_{k=1}^K \sum_{k'=1}^K \beta_{kk'} LnX_k LnX_{k'} + \frac{1}{2} \sum_{l=1}^{L-1} \sum_{l'=1}^{L-1} \alpha_{ll'} Ln(Y_l / Y_L) Ln(Y_{l'} / Y_L) + \sum_{k=1}^K \sum_{l=1}^{L-1} \xi_{kl} LnX_k Ln(Y_l / Y_L) + U \quad (8)$$

Stochastinė ribos analizė

Pirmieji stochastinės ribinės našumo funkcijos modelį pasiūlė Aigner, Lovell ir Schmidt (1977). Kadangi šis pagrindinis modelis buvo plačiai naudojamas, o mūsų tikslai yra sutelkti į atstumo funkciją vietoj prielaidų nustatymo regresijoms, ši analizė yra ribojama iki pusės normalės neefektyvumo komponentų nustatymo.

Modelį galima apibrėžti taip:

$$y_i = \beta x_i + V_i - U_i, \quad i = 1, 2, \dots, N, \quad (10)$$

Kur y yra išiega; x yra sąnaudų vektorius; β yra parametru vektorius. Stochastinė riba yra nustatoma išskaidant klaidos terminą į dvi sudėtines dalis, techninį neefektyvumą (U_i) ir atsitiktinę klaidą (V_i). Pagrindinė stochastinė riba teigia, kad U_i yra pusė normalės ir nepriklauso nuo V_i . Atsitiktinės klaidos laikomos normalėmis. *Monte Karlo analizės kūrimas*. Šiame eksperimente, mūsų Monte Karlo analizei nustatyta 1000 atsakymų. Į pavyzdį įeina $\alpha, \beta, \xi, \gamma, \sigma, N$ ir X , kur N yra pavyzdžio dydis ir X yra sąnaudų vektorius. Mes priverčiame pavyzdžio dydį keistis nuo 50 iki 800. Iš Tim Coelli (1995) gautų rezultatų matyti, kad γ paklaida mažėja, kai γ vertė artėja prie vieneto. Kadangi mūsų tyrimas nagrinėja tik atstumo funkcijas, vietoj stochastinių ribos regresijų, mes nustatome γ reikšmę ties 0.96 ir σ ties 0.5. Mes sumažiname 'X'-ų skaičių tik iki dviejų kintamųjų, vienas yra pastovus, o kitas yra sąnaudos, sukurtos iš pastovaus (5,15). Mes sukuriame išiega iš funkcijos(6) vietoj įprastinės funkcijos(10), nes galutinė stochastinė ribinė atstumo funkcija(8) pasikeičia iš *translog* atstumo funkcijos(6) ir todėl įprastinė stochastinė ribos funkcija(10) negali duoti daugkartinės išiegos mūsų eksperimentuose. Norėdami patikrinti daugkartinės išiegos atstumo funkciją, mes atliksime abu eksperimentus - su 2 išiegom ir su 3 išiegom.

Funkcijoje (6), mes taip pat turime pateikti techninio neefektyvumo komponentus $LnDi(x, y)$ ir atsitiktinės klaidos komponentus V_i . Kiekviename atsakyme du standartinių normalių atsitiktinių skaičių, kurių ilgis 800, vektoriai buvo sukurti siekiant sukurti techninio neefektyvumo komponentus ir atsitiktinės klaidos komponentus. Čia, V_i yra sukurtas dauginant pirmą atsitiktinių skaičių iš σ_V ir U_i yra suformuotas dauginant absoliučias antrojo atsitiktinių skaičių vektoriaus reikšmes iš σ_U . Reikėtų pabrėžti, kad σ_V ir σ_U buvo gauti iš γ ir σ . Per funkcijas $\sigma^2 = \sigma_U^2 + \sigma_V^2$ ir $\gamma = \sigma_U^2 / \sigma^2$, galime sužinoti, kad σ_V yra fiksuotas ties 0.1 ir σ_U yra $\sqrt{0.24}$. Remiantis funkcija (6), mes manome, kad visos išiegos yra teigiamos. Ši prielaida sudaro papildomų klūčių išiegos vertei, gautai iš funkcijos (6). Dar daugiau- mes galime atlikti skirtingus eksperimentus skirtingoms išiegos kombinacijoms. Tokiu būdu, mes turime tris scenarijus atskirai esant vienodai išiegai, 2 išiegom ir 3 išiegom.

Monte Karlo eksperimentų rezultatai. Iš funkcijos (12) aišku, kad vidutinis techninis našumas TE (*plg. angl. technical efficiency*) bus nulemtas ir γ , ir σ . Coelli (1995) teigė, kad paklaida γ įvertinime duos klaidinančią informaciją apie vidutinį TE lygį. Teigiama paklaida γ įvertinime, kai jo tikroji reikšmė yra nulis, duoda neigiamą paklaidą TE įvertinime. O σ paklaida padarys įtaką TE įvertinimų paklaidos dydžiui. Taigi, jei γ ir σ yra objektyvūs, tada vidutinio techninio našumo įvertinimas taip pat yra objektyvus. Todėl, šiame tyrime mes tik trumpai apžvelgiame γ ir σ įvertinimų paklaidas ir MSE (*plg. angl. Mean Squared Error* – vidutinė kvadratinė klaida) iš Monte Karlo eksperimentų rezultatų. γ ir σ įvertinimų paklaida ir vidutinė kvadratinė klaida parodyti 1, 2, 3 ir 4 paveiksluose. Ir γ , ir σ absoliučios paklaidų ir MSE reikšmės, esant vienodos išiegos scenarijui yra mažiausios iš trijų scenarijų. Tai taip pat parodė, kad vienodos išiegos scenarijuje, γ paklaida ir MSE visada yra žemesnės už σ paklaidą ir MSE. Tačiau, jie visi priartėja prie 0, kai pavyzdžio dydis yra 800. Šis rezultatas atitinka išvadą, kurią padarė Coelli (1995). Jis teigia, kad paklaida mažėja iki nulio, kai γ priartėja prie vieneto ir pavyzdžio dydis padidėja iki 800. (vienodos išiegos scenarijaus įvertinimo paklaida mūsų tyrime yra teigiama, bet paklaidos su viena išiega eksperimentuose Coelli (1995) tyrime yra neigiamos, kai γ priartėja prie vieneto. Priežastis yra ta, kad vienodos išiegos funkcija (8) yra išvesta iš atstumo funkcijos. Ir ji turi neigiamą priklausomo kintamojo $Ln(Y_L)$ ženklą.). γ ir σ absoliučios paklaidų ir MSE reikšmės esant 2-šiegių scenarijui yra visos aukštesnės, negu esant vienodos išiegos scenarijui ir žemesnės, negu esant 3-šiegių scenarijui. Įvertinimo paklaidos esant 3-šiegių scenarijui yra neigiamos ir absoliučios jų reikšmės yra didžiausios iš trijų scenarijų. Visi anksčiau pateikti rezultatai parodo, kad stochastinės ribinės atstumo funkcijos modelis su keliomis išiegomis tikriausiai turi mažiau įvertinimų su paklaidomis. Daugiau išiegu reiškia daugiau paklaidų įvertinimuose ir įvertintame vidutiniame techniniame našume.

Išvada. Stochastinė ribinė atstumo funkcija, išskyrus scenarijų su vienoda išiega, nusileis paklaidos įvertinimams net kai pavyzdys yra didelio dydžio. 2-šiegių modelis duos geresnius įvertinimus negu 3-šiegių modelis. Didėjantis pavyzdžio dydis pagerins santykinę stochastinės ribinės atstumo funkcijos ML įvertinimų veiklą. Taigi Monte Karlo analizė pagaliau patvirtino rezultatą, kad stochastinė ribinė atstumo funkcija tikriausiai yra daugkartinės išiegos našumo paklaida. Tai naujas ir svarbus metodo taikymo rezultatas su šiuo tyrimu susijusioje srityje.

Raktažodžiai: *Stochastinė ribinė atstumo funkcija; Monte Karlo analizė; paklaida; įvertinimas.*

The article has been reviewed.

Received in June, 2011; accepted in June, 2012.