World Applied Sciences Journal 13 (8): 1864-1869, 2011
ISSN 1818-4952

Malmquist Productivity Index with Stochastic Variations in Data

M.A. Raayatpanah and H. Ghasvari

Kashan Branch, Islamic Azad University, Kashan, Iran

Abstract: The performance of a decision making unit(DMU) can be evaluated in either across-sectional or a time-series manner and data envelopment analysis (DEA) is a useful method for both types of evaluation. Productivity growth is decomposed using a generalized Malmquist productivity index. In response to the criticism that in most applications there is error and random noise in the data, a number of mathematically elegant solutions to incorporating stochastic variations in data have been proposed. In this paper, we study the Malmquist productivity index, that are the result of a stochastic process.

Key words: Data envelopment analysis · Malmquist productivity index · Stochastic data envelopment analysis (SDEA)

INTRODUCTION

Data envelopment analysis (DEA) is a mathematical programming technique, which is used to evaluate the relative efficiency of decision making units (DMUs) and has been proposed by Charnes *et al.* [1] as the CCR model, (the model by Banker *et al.* [2] is usually referred to as the BCC model). The original idea behind DEA was to provide a methodology whereby, within a set of comparable decision making units (DMUs), those exhibiting best practice could be identified and would form an efficient frontier. Furthermore, the methodology enables one to measure the level of efficiency of non-frontier units and to identify benchmarks against which such inefficient units can be compared. The purpose of the current paper is to using of the DEA for evaluating the performance of multiple comparable Queueing. The performance is inclusive identify efficiency, benchmarks, returns to scale, ranking, sensitivity analysis.

Performance measurement is an important issue for at least two reasons. One is that in a group of units where only limited number of candidates can be selected, the performance of each must be evaluated in a fair and consistent manner. The other is that as time progresses, better performance is expected. Hence, the units with declining performance must be identified in order to make the necessary improvements. Hereafter, a unit to be evaluated is referred to as a decision making unit (DMU).

In addition to comparing the relative performance of a set of DMUs at a specific period, the conventional DEA can also be used to calculate the productivity change of a DMU over time. Caves *et al.* [3,4] proposed a Malmquist productivity index (MPI) which calculates the relative performance of a DMU at different periods of time using the technology of a base period.Since the base period used to define the production technology affects the results, several modifications for calculating MPI have been proposed. The most popular method is the one proposed by Färe et al. [5] which takes the geometric mean of the MPIs calculated from two base periods. Later, Pastor and Lovell [6] proposed a global MPI, based on a technology defined by DMUs of all periods, to calculate productivity changes.

An early criticism of DEA has been that it assumes data to be deterministic. A distinction has been made in the literature in that DEA-type approaches yield efficiency measures, while statistical approaches (stochastic frontier models) produce efficiency estimates (Horace and Schmidt [7]). In other words, the DEA approach has been deemed non-statistical [8, 9]. The many and varied responses to this criticism have followed Timmer [10] in introducing noise in the input and output constraints. One of the earliest of these responses involved the development of chance constrained formulations of the mathematical programs underlying the DEA problem in order to accommodate stochastic variations in data, e.g., [11-19].

Corresponding Author: M.A. Raayatpanah, Kashan Branch, Islamic Azad University, Kashan, Iran.
E-mail: raayatpanah@khayam.ut.ac.ir.

In this paper, we offer using Data Envelopment Analysis for evaluating the performance of Malmquist productivity index for DMUs, that are the result of a stochastic process.

The remainder of this paper has the following structure: in section 2, we present the required background. Section 3 introduces our method as a usage of Malmquist productivity index, that are the result of a stochastic process. Section 4 illustrates the proposed method using an example. Finally, conclusions are given in section 5.

Background

DEA Models: Data envelopment analysis (DEA) is a method for evaluating efficiency of decision making units (DMUs). Consider n decision making units DMU_j ($j = 1,2,...,n$), each DMU_j consuming input levels x_{ij} ($i = 1,2,...,n$) to produce output levels y_{rj} ($r = 1,2,...,s$). The relative efficiency score of DMU_o under the CCR model is given by the following optimization problem:

$$\begin{aligned} &Max\ \frac{u^T y_o}{v^T x_o} \\ &s.t.\ \frac{u^T y_j}{v^T x_j} \le 1\ j = 1,2,\ldots,n \\ &u,v \ge 0 \end{aligned} \tag{2}$$

Where u and v represent vectors for the output and input weights, respectively.

We point out that the DEA model (1) is equivalent to the following linear program which is called the output-oriented formulation for the CCR model:

$$\begin{aligned} &Min\ v^T x_o \\ &s.t.\ u^T y_o = 1 \\ &-v^T X + u^T Y \le 0 \\ &u,v \ge 0 \end{aligned} \tag{2}$$

Also, problem (1) can be converted to the following linear program (LP), which is essentially the CCR model in input-oriented and envelopment form:

$$\begin{aligned} &Min\ \theta \\ &s.t.\ \sum_{j=1}^{n} \lambda_j x_j \le \theta x_o \\ &\sum_{j=1}^{n} \lambda_j y_j \ge y_o \end{aligned} \tag{3}$$

Malmquist Productivity Index: We assume that for each time period $t = 1,...,T$, the production technology S^t models the transformation of inputs, x^t, into outputs, y^t, $S^t = \{(x^t,y^t): x^t canproducey^t\}$. Now, we the DEA score θ of the period r DMU_o measured by means of the period k frontier, we denote it as $D_o^k(x^l,y^l)$. Then, we have:

$$\begin{aligned} &D_o^k(x^l, y^l) = Min \qquad \theta \\ &s.t.\ \sum_{j=1}^{n} \lambda_j x_j^k \le \theta x_o^l \\ &\sum_{j=1}^{n} \lambda_j y_j^k \ge y_o^l \\ &\lambda_j \ge 0,\ j = 1,...,n. \end{aligned} \tag{4}$$

that, $k,l \in \{t,t+1\}$

Malmquist productivity index was illustrated by Caves *et al.* [3, 4] and listed as follows:

$$M^t = \frac{D_o^t(x^{t+1}, y^{t+1})}{D_o^t(x^t, y^t)} \tag{5}$$

In this formulation, technology in period t is the reference technology. The follow equation represents the productivity of the production point (x^{t+1},y^{t+1}) relative to the production point (x^t,y^t). A value >1 will indicate positive TFP growth from period t to period $t + 1$ and vice versa.

$$TFP_O(x^{t+1}, y^{t+1}, x^t, y^t) = \sqrt{\frac{D_O^t(x^{t+1}, y^{t+1})}{D_O^t(x^t, y^t)} \times \frac{D_O^{t+1}(x^{t+1}, y^{t+1})}{D_O^{t+1}(x^t, y^t)}} \tag{6}$$

In the assumption of CRS, the above index can be broken down in to technological change (TECH) and technical efficiency change (EFFCH) indexes The equation can be written as.

$$TFP_O(x^{t+1}, y^{t+1}, x^t, y^t) = \frac{D_O^t(x^{t+1}, y^{t+1})}{D_O^t(x^t, y^t)} \times \sqrt{\frac{D_O^t(x^{t+1}, y^{t+1})}{D_O^{t+1}(x^{t+1}, y^{t+1})} \times \frac{D_O^t(x^t, y^t)}{D_O^{t+1}(x^t, y^t)}} \tag{7}$$

TFP_o measures the productivity change between periods t and $t + 1$. Productivity declines if $TFP < 1$, remains unchanged if $TFP = 1$ and improves if $TFP > 1$.

Stochastic DEA: We follow the notation in Cooper *et al.* [12] and let $\tilde{x}_j = (\tilde{x}_{1j},\ldots,\tilde{x}_{mj})^T$ and $\tilde{y}_j = \tilde{y}_{1j},\ldots,\tilde{y}_{sj})^T$ represent $(m \times 1)$ and $(s \times 1)$ random input and output vectors and $x_j = (x_{1j},\ldots,x_{mj})^T$ and $y_j = (y_{1j},\ldots,y_{sj})^T$ stand for the corresponding vectors of expected values of input and output for each $DMU_j, j = 1,\ldots,n$. That is, we utilize these expected values in place of the observed values in (3). See Olesen and Petersen [17] for an alternate approach which uses the means of a series of observations to obtain confidence interval estimates.

Let us consider all input and output components to be jointly normally distributed in the following chance constrained version of a stochastic DEA model:

$$
\begin{aligned}
&Min \quad \theta \\
&s.t. \; Pr\{\sum_{j=1}^{n}\lambda_j\tilde{x}_{ij} \le \theta\tilde{x}_{io}\} \ge 1-\alpha, \; i = 1,\ldots,m, \\
&Pr\{\sum_{j=1}^{n}\lambda_j\tilde{y}_{rj} \ge \tilde{y}_{ro}\} \ge 1-\alpha, \; r = 1,\ldots,s, \qquad (8) \\
&\lambda_j \ge 0 \; j = 1,\ldots,n.
\end{aligned}
$$

Here, Pr means Probability and α is a predetermined number between 0 and 1. We now use this model to define stochastic efficiency as follows.

Definition. (Stochastic Efficiency) DMU_o is stochastic efficient if and only if for the optimal solutions $\theta^* = 1$.

Now, we show how to obtain the θ^*, from deterministic equivalents of the stochastic models represented in (8). With normal distributions and zero order decision rules we can obtain a deterministic equivalent for (8) which can be represented by.

$$
\begin{aligned}
&Min \quad \theta \\
&s.t. \; \sum_{j=1}^{n}\lambda_j x_{ij} - \Phi^{-1}(\alpha)\sigma_i^I(\theta,\lambda) \le \theta x_{io}, \; i = 1,\ldots,m, \\
&\sum_{j=1}^{n}\lambda_j y_{rj} - \Phi^{-1}(\alpha)\sigma_r^O(\lambda) \ge y_{ro}, \; r = 1,\ldots,s, \qquad (9) \\
&\lambda_j \ge 0 \; j = 1,\ldots,n.
\end{aligned}
$$

Where the x_{ij} and y_{rj} (including x_{io} and y_{ro}) are the means of these variables. (these means are assumed to be known). Here Φ is the standard normal distribution function and Φ^{-1}, its inverse, Finally,

$$
(\sigma_i^I(\theta,\lambda))^2 = \sum_{j\ne o}\sum_{k\ne o}\lambda_j\lambda_k Cov(\tilde{x}_{ij},\tilde{x}_{ik}) + 2(\lambda_o - \theta)
$$

$$
\sum_{j\ne o}\lambda_j Cov(\tilde{x}_{ij},\tilde{x}_{io}) + (\lambda_o - \theta)^2 Var(\tilde{x}_{io}),
$$

and

$$
(\sigma_r^O(\lambda))^2 = \sum_{i\ne o}\sum_{j\ne o}\lambda_i\lambda_j Cov(\tilde{y}_{ri},\tilde{y}_{rj}) + 2(\lambda_o - 1)
$$

$$
\sum_{i\ne o}\lambda_i Cov(\tilde{y}_{ri},\tilde{y}_{ro}) + (\lambda_o - 1)^2 Var(\tilde{y}_{ro}),
$$

Where we have separated out the terms for DMU_o. Thus, θ^* can be determined from (9) where the data (means and variances) are all assumed to be known.

Because of the functional forms of $\sigma_i^I(\theta,\lambda)$ and $\sigma_r^O(\lambda)$ it is obvious that models (9) is non-linear programming problem. Let ω_i^I and ω_r^O be non-negative variables. Replace $\sigma_i^I(\theta,\lambda)$ by ω_i^I and $\sigma_r^O(\lambda)$ by ω_r^O in (9) and add two quadratic equality constraints, $(\omega_i^I)^2 = (\sigma_i^I(\theta,\lambda))^2$ and $(\omega_r^O)^2 = (\sigma_r^O(\lambda))^2$, to (9), then (9) is transformed to easily solvable quadratic programming problems.

To simplify matters in a different manner let us assume that only DMU_o has random variations in its inputs and outputs, i.e., $\sigma_{io}^I \ne 0$, $\sigma_{ro}^O \ne 0$, $\sigma_{ij}^I = 0$ and $\sigma_{rj}^O = 0$ $(j \ne 0)$ for all i and r. In this case, model (9) can be written.

$$
\begin{aligned}
&Min \quad \theta \\
&s.t. \; \sum_{j=1}^{n}\lambda_j x'_{ij} \le \theta x'_{io}, \; i = 1,\ldots,m, \\
&\sum_{j=1}^{n}\lambda_j y'_{rj} \ge y'_{ro}, \; r = 1,\ldots,s, \qquad (10) \\
&\lambda_j \ge 0 \; j = 1,\ldots,n.
\end{aligned}
$$

Where

$$
x'_{io} = x_{io} - \Phi^{-1}(\alpha)\sigma_{io}^I, \; i = 1,\ldots,m
$$

$$
x'_{ij} = x_{ij}, \; j \ne o, i = 1,\ldots,m,
$$

$$
y'_{ro} = y_{ro} - \Phi^{-1}(\alpha)\sigma_{ro}^O, \; r = 1,\ldots,s
$$

$$
y'_{rj} = y_{rj}, \; j \ne o, r = 1,\ldots,s,
$$

With these assumptions model (10) is the deterministic equivalent of stochastic model (8).

Malmquist Productivity Index with Stochastic Data: We assume that for each time period $t = 1,..,T$ and let $\tilde{x}_j^t = (\tilde{x}_{1j}^t,\ldots,\tilde{x}_{mj}^t)^T$ and $\tilde{y}_j^t = \tilde{y}_{1j}^t,\ldots,\tilde{y}_{sj}^t)^T$ represent random input and output vectors and $x_j^t = (x_{1j}^t,\ldots,x_{mj}^t)^T$ and $y_j^t = (y_{1j}^t,\ldots,y_{sj}^t)^T$ stand for the corresponding vectors of expected values of input and output for each DMU_j, $j = 1,\ldots,n$ of the period t. Now, we the DEA score θ of the period l DMU_o measured by means of the period k frontier, we denote it as $D_O^k(\tilde{x}^l,\tilde{y}^l)$. Then, we have:

$$D_o^k(\tilde{x}^l,\tilde{y}^l) = Min\ \ \theta$$

$$s.t.\ \sum_{j=1}^{n}\lambda_j x_{ij}^k - \Phi_k^{-1}(\alpha)\sigma_i^{I,k}(\theta,\lambda) \le \theta x_{io}^l,\ i = 1,\ldots,m,$$

$$\sum_{j=1}^{n}\lambda_j y_{rj}^k - \Phi_k^{-1}(\alpha)\sigma_r^{O,k}(\lambda) \ge y_{ro}^l,\ r = 1,\ldots,s, \tag{13}$$

$$\lambda_j \ge 0,\ \ j = 1,\ldots,n.$$

that, k, $l \in \{t,t+1\}$. Here, Φ_k is the standard normal distribution function of the period k and Φ_k^{-1}, its inverse, Finally,

$$(\sigma_i^{I,k}(\theta,\lambda))^2 = \sum_{j\ne o}\sum_{p\ne o}\lambda_j\lambda_p Cov(\tilde{x}_{ij}^k,\tilde{x}_{ip}^k) + 2(\lambda_o - \theta)$$

$$\sum_{j\ne o}\lambda_j Cov(\tilde{x}_{ij}^k,\tilde{x}_{io}^k) + (\lambda_o - \theta)^2 Var(\tilde{x}_{io}^k),$$

and

$$(\sigma_r^{O,k}(\lambda))^2 = \sum_{i\ne o}\sum_{j\ne o}\lambda_i\lambda_j Cov(\tilde{y}_{ri}^k,\tilde{y}_{rj}^k) + 2(\lambda_o - 1)$$

$$\sum_{i\ne o}\lambda_i Cov(\tilde{y}_{ri}^k,\tilde{y}_{ro}^k) + (\lambda_o - 1)^2 Var(\tilde{y}_{ro}^k),$$

Which, in section (2.3) have described the computational scheme.

Then, Malmquist productivity index listed as follows:

$$M^t = \frac{D_o^t(\tilde{x}^{t+1},\tilde{y}^{t+1})}{D_o^t(\tilde{x}^t,\tilde{y}^t)} \tag{14}$$

and

$$TFP_O(\tilde{x}^{t+1},\tilde{y}^{t+1},\tilde{x}^t,\tilde{y}^t) = \sqrt{\frac{D_O^t(\tilde{x}^{t+1},\tilde{y}^{t+1})}{D_O^t(\tilde{x}^t,\tilde{y}^t)} \times \frac{D_O^{t+1}(\tilde{x}^{t+1},\tilde{y}^{t+1})}{D_O^{t+1}(\tilde{x}^t,\tilde{y}^t)}} =$$

$$\frac{D_O^t(\tilde{x}^{t+1},\tilde{y}^{t+1})}{D_O^t(\tilde{x}^t,\tilde{y}^t)} \times \sqrt{\frac{D_O^t(\tilde{x}^{t+1},\tilde{y}^{t+1})}{D_O^{t+1}(\tilde{x}^{t+1},\tilde{y}^{t+1})} \times \frac{D_O^t(\tilde{x}^t,\tilde{y}^t)}{D_O^{t+1}(\tilde{x}^t,\tilde{y}^t)}} \tag{15}$$

TFP measures the productivity change between periods t and $t + 1$. Productivity declines if *TEP* < 1, remains unchanged if *TFP* = 1 and improves if *TFP* > 1.

Numerical Example: Here, we present one example. Suppose we have two time periods $t = 1,2$ and Consider six decision making units DMU_j ($j = 1,2,..,6$), each DMU_j consuming two random input vectors $\tilde{x}_{ij}^t (i=1,2)$ to produce two random output vectors $\tilde{y}_{rj}^t (r=1,2)$.

Table 1 presents means, variances for the input and output vectors of DMUs in two time periods $t = 1,2$.

Based on (9-10) and (13-14), the adjusted input and output for DMUs is presented and the results of Malmquist productivity index with stochastic data are shown in Table 2, which are the adjusted score explained in Section 3 for $\alpha = 0.05$ and $\alpha = 0.01$.

Regarding the performance improvement, the values in the last columns of Table 2 present that for $\alpha = 0.05$ two DMUs, A and C, have *TFP* > 1 that means they declines. DMU D, has *TFP* = 1 that means it remains unchanged. The average *TFP* of the six DMUs is 0.958. Hence, in general, the performances of the six DMUs have declined after the reorganization.

Table 1: The means, variances for the random input and output vectors of DMUs in two time periods $t = 1,2$.

DMU	input 1	input 2	output 1	output 2
A	4.5 (1.2)	3.2 (1.2)	3.5 (1.6)	4.5 (1.2)
	3.4 (2.5)	2.3 (0.5)	2.4 (1.5)	3.8 (1.5)
B	3.5 (1.2)	3.8 (1.3)	3.7 (1.2)	2.5 (0.4)
	4.4 (2.5)	3.4 (1.5)	3.4 (1.5)	3.6 (1.7)
C	1.4 (0.4)	5.5 (2.4)	4.5 (1.6)	3.4 (1.6)
	5.3 (1.5)	6.4 (2.3)	5.2 (2.5)	3.6 (1.5)
D	3.3 (1.5)	4.5 (1.5)	4.3 (1.4)	3.5 (1.2)
	3.5 (2.6)	4.5 (2.3)	5.2 (1.6)	5.4 (2.3)
E	4.3 (1.2)	6.5 (1.8)	3.8 (1.4)	4.7 (1.3)
	5.3 (1.2)	4.7 (1.9)	5.3 (2.1)	6.4 (2.3)
F	5.3 (1.6)	3.6 (1.8)	4.5 (1.6)	5.4 (1.8)
	5.3 (1.6)	6.4 (2.5)	5.4 (1.5)	5.4 (2.3)

.: means, (.): variances, $\frac{t=1}{t=2}$

Table 2: The results of Malmquist productivity index with stochastic data for $\alpha = 0.05$ and $\alpha = 0.1$

DMU	$\alpha = 0.05$	$\alpha = 0.1$	Result for α 0.05	Result for α 0.1
A	1.20	1.10	Improves	Improves
B	0.90	0.85	Declines	Declines
C	1.10	1.10	Improves	Improves
D	1.00	0.95	Unchanged	Declines
E	0.80	1.00	Declines	Unchanged
F	0.75	0.70	Declines	Declines

CONCLUSIONS

Malmquist Productivity Index (MPI), based on DEA, is used to measure the performance changes over time. The Malmquist Productivity Index allows us to distinguish between shifts in the production frontier (technological change) and movement of departments nearer the frontier (efficiency change).

However, since DEA does not account for statistical noise, estimates of efficiency will be biased when stochastic elements are a prominent feature of the true production process or the variables used in the analysis are measured with error. So, in this paper, we study the Malmquist productivity index, that are the result of a stochastic process.

ACKNOWLEDGMENT

This work was supported in part by Islamic Azad University Kashan beranch and M.A. Raayatpanah Authoer would like to express his appreciationto his sponsor.

REFERENCES

1. Charnes, A., W.W. Cooper and E. Rhodes, 1978. Measuring the efficiency of decision making units. European J. Operational Res., 2: 429-444.
2. Banker, R.D., A. Charnes and W.W. Cooper, 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Sci., 30(9): 1078-1092.
3. Caves, D.W., L.R. Christensen and W.E. Diewert, 1982. The economic theory of index numbers and the measurement of input, output and productivity. Econo-metrica, 50: 1393-1414.
4. Caves, D.W., L.R. Christensen and W.E. Diewert, 1982. Multilateral comparisons of output, input and productivity using superlative index numbers. The Economic J., 92: 73-86.
5. Färe, R., G. rosskopf, M. Norris and Z. Zhang, 1994. Productivity growth, technical progress and efficiency change in industrialized countries. American Economic Review, 84: 66-83.
6. Pastor, J.T. and C.A.K. Lovell, 2005. A global Malmquist productivity index. Economics Letters, 88: 266-271.
7. Horace, W.C. and P. Schmidt, 1996. Confidence statements for efficiency estimates from stochastic frontier models. J. Productivity Analysis, 7: 257-282.
8. Banker, R.D., 1996. Hypothesis test using data envelopment analysis. J. Productivity Analysis, 7(1): 39-59.
9. Gong, B.H. and R.C. Sickles, 1992. Finite sample evidence on the performance of stochastic frontiers and data envelopment analysis using pane: data. J. Econometrics, 51: 27-56.
10. Timmer, C.P., 1971. Using probabilistic frontier production function to measure technical efficiency. J. Political Economy, 79: 776-794.
11. Banker, R.D., 1986. Stochastic data envelopment analysis. Pittsburgh: Carnegie-Mellon University.
12. Cooper, W.W., Z.M. Huang and S.X. Li, 1996. Satisficing DEA models under chance constraints. Annals of Operations Res., 66: 279-295.
13. Cooper, W.W., Z.M. Huang, V. Lelas, S.X. Li and O.B. Olesen, 1998. Chance constrained programming formulations for stochastic characterisations of efficiency and dominance in DEA. J. Productivity Analysis, 9: 53-79.
14. Desai, 1986. Extensions to measures of relative efficiency with an application to educational productivity. PhD dissertation, University of Pennsylvania, Philadelphia.
15. Land, K.C., C.A.K. Lovell and S. Thore, 1993. Chance constrained data envelopment analysis. Managerial and Decision Economics, 14: 541-554.

16. Morita, H. and L. Seiford, 1999. Characteristics on stochastic DEA efficiency. J. the Operations Research Society of Japan, 42: 389-404.
17. Olesen, O. and N.C. Petersen, 1995. Chance constrained efficiency evaluation. Management Sci., 41: 442-457.
18. Sengupta, J.K., 1987. Data envelopment analysis for efficiency measurement in the stochastic case. Computers and Operations Res., 14: 117-129.
19. Sueyoshi, T., 2000. Stochastic DEA for restructure strategy: an application to a Japanese petroleum company. Omega, 28: 385-398.