Applied Mathematical Sciences, Vol. 2, 2008, no. 9, 443 - 448

IDEA-Based Malmquist Productivity Index

Gh. Tohidi [a, 1], Sh. Razavyan [b] and F. Baniasadi Rad [a]

[a] Department of Mathematics, Islamic Azad University
Tehran-Central Branch, Tehran, Iran

[b] Department of Mathematics, Islamic Azad University
Tehran-South Branch, Tehran, Iran

Abstract

Data Envelopment Analysis (DEA) based Malmquist productivity index measures the productivity change over time. The current paper defines Imprecise DEA (IDEA) based Malmquist productivity index in order to measure the productivity change of DMU, with imprecise data.

Keywords: DEA, Imprecise Data Envelopment Analysis (IDEA), Malmquist Index

1. Introduction

Data Envelopment Analysis (DEA) provides a suitable way to estimate the relative efficiency of Decision Making Units (DMUs). DEA assumes that the value for all inputs and outputs be known exactly. This assumption may not be true. For example, some outputs and inputs may be only known as in forms of bounded data, ordinal data, and ratio bounded data. In this case, the resulting DEA model is a non - linear program, and is called imprecise DEA(IDEA). There are two different approaches in dealing with imprecise outputs and inputs. Cooper et al. [4] and Kim et al. [7] show that by using scale transformations and variable alternations, the non-linear IDEA can be transformed into a linear program. In a recent study of Chen et al. [3] and Zhu [10] the non-linear IDEA is solved in the standard linear CCR model [2] via identifying a set of exact data from the imprecise data when only bounded data and weak ordinal data are present. Zhu [11] extends the method of Chen et al. [3] and Zhu [10] to deal with strong ordinal data and ratio bounded data. The Malmquist index was introduced by Caves et al. [1], who dubbed it the Malmquist productivity index after Sten Malmquist [8]. Fare et al.

[1]Corresponding author, P. O. Box 14515-459, e-mail: ghatohidi@yahoo.com

[5] construct the DEA-based Malmquist productivity index as the geometric mean of two Malmquist productivity indexes of Caves et al. [1], which are based on the Shepherd's [9] distance functions. Fare et al. [6] decomposed productivity change into efficiency change and technical change components and used non-parametric linear programming models for its computation. In this paper IDEA-based Malmquist productivity index is proposed. The rest of the paper is organized as follows. Section 2 presents the imprecise data and IDEA. IDEA-based Malmquist productivity is presented in section3. Section 4 concludes.

2. Imprecise Data and IDEA

Imprecise data are in the forms of bounded data, ordinal data, and ratio bounded data.
i) Bounded data is defined by use of lower and upper bound as follows:

$$\underline{y}_{rj} \leq y_{rj} \leq \overline{y}_{rj}, \quad r \in BO, \quad \underline{x}_{ij} \leq x_{ij} \leq \overline{x}_{ij}, \quad i \in BI, \tag{1}$$

where $\underline{y}_{rj}$ and $\underline{x}_{ij}$ are the lower bounds and $\overline{y}_{rj}$ and $\overline{x}_{ij}$ are the upper bounds, and BO and BI represent the associated sets the containing bounded outputs and bound inputs, respectively.
ii) Weak ordinal data are as follows:

$$y_{rj} \leq y_{rk}, j \neq k, r \in DO, \quad x_{ij} \leq x_{ik}, j \neq k, i \in DI, \tag{2}$$

or, to simplify the presentation,

$$y_{r1} \leq \cdots \leq y_{rk} \leq \cdots \leq y_{rn}, r \in DO, \; x_{i1} \leq \cdots \leq x_{ik} \leq \cdots \leq x_{in}, i \in DI, \tag{3}$$

where DO and DI represent the associated sets containing weak ordinal outputs and inputs, respectively. Strong ordinal data are as follows:

$$y_{r1} < \cdots < y_{rk} < \cdots < y_{rn}, r \in SO, \; x_{i1} < \cdots < x_{ik} < \cdots < x_{in}, i \in SI, \tag{4}$$

where SO and SI represent the associated sets containing strong ordinal outputs and inputs, respectively.
iii) Ratio bounded data are as follows:

$$L_{rj} \leq \frac{y_{rj}}{y_{rjo}} \leq U_{rj}, j \neq j_o, r \in RO, \; G_{ij} \leq \frac{x_{ij}}{x_{ijo}} \leq H_{ij}, j \neq j_o, i \in RI, \tag{5}$$

where G_{ij} and L_{rj} represent the lower bounds, and U_{rj} and H_{ij} represent the upper bounds, and RO and RI represent the associated sets containing ratio bounded outputs and inputs, respectively. If we incorporate (1)-(5) into CCR model [2], we have:

$$\begin{array}{lll} \text{Max} & \sum_{r=1}^{s} u_r y_{ro} & \\ s.t. & \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, & j = 1, \ldots, n, \\ & \sum_{i=1}^{m} v_i x_{io} = 1, & \\ & x_i = (x_{ij}) \in D_i^-, & i = 1, \ldots, m, \\ & y_r = (y_{rj}) \in D_r^+, & r = 1, \ldots, s, \\ & v_i \geq 0, u_r \geq 0, & i = 1, \ldots, m, r = 1, \ldots, s, \end{array} \tag{6}$$

where $(x_{ij}) \in D_i^-$ and $(y_{rj}) \in D_r^+$ represent any of or all of (1)-(4). Obviously, model (6) is non-linear, because some of the outputs and inputs become unknown decision variables. The following theorem provides the theoretical foundation to the approach developed by Chen et al. [3] and Zhu [10].

Theorem 1. *Suppose D_r^+ and D_i^- are given by (1). Then for DMU_o the optimal value to (6) can be achieved at $y_{ro} = \overline{y}_{ro}$ and $x_{io} = \underline{x}_{io}$ for DMU_o and $y_{rj} = \underline{y}_{rj}$ and $x_{ij} = \overline{x}_{ij}$ for $DMU_j, j \neq o$.*
Proof: See Zhu [11]. □

Chen et al. [3] and Zhu [10] identify a set of exact data from bounded data and weak ordinal data, and then use the linear CCR model [2] to solve the non-linear IDEA model (6). Theorem 1 shows that when DMU$_o$ is under evaluation, we can have a set of exact data via setting $y_{ro} = \overline{y}_{ro}$ and $x_{io} = \underline{x}_{io}$ for DMU$_o$ and $y_{rj} = \underline{y}_{rj}$ and $x_{ij} = \overline{x}_{ij}$ for DMU$_j, j \neq o$, while model (6) maintains the efficiency rating for DMU$_o$. Note that in this case, model (6) is no longer a non-linear program, but a (linear) CCR model [2] and its dual is as follows:

$$\begin{array}{lll} \text{Min} & \theta_o & \\ s.t. & \sum_{j=1, j\neq o}^{n} \lambda_j \overline{x}_{ij} + \lambda_o \underline{x}_{io} \leq \theta_o \underline{x}_{io}, & i \in BI, \\ & \sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta_o x_{ij}, & i \notin BI, \\ & \sum_{j=1, j\neq o}^{n} \lambda_j \underline{y}_{rj} + \lambda_o \overline{y}_{ro} \geq \overline{y}_{ro}, & r \in BO, \\ & \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro}, & r \notin BO \\ & \lambda_j \geq 0, & j = 1, \ldots, n, \end{array} \tag{7}$$

where,y_{rj}, $(r \notin BO)$ and $x_{ij}, (i \notin BI)$ are exact data . Model (7) is used for calculating Malmquist productivity index as is described in next section.

3.IDEA-Based Malmquist Productivity Index

IDEA-based Malmquist productivity index measures the productivity change of decision making units with imprecise data to define the IDEA-based Malmquist

index of productivity change. Suppose the following panel of $i = (1, \ldots, I)$ production processes observed in $t = (1, \ldots, T)$ periods in transformation a series of input vectors $x_i^t = (x_{i1}^t, \ldots, x_{im}^t) \in R_+^m$ into output vector $y_i^t = (y_{i1}^t, \ldots, y_{is}^t) \in R_+^s$. Note that some outputs and input may be imprecise. Production technology can be representing as can produce . Thus it is possible to express the benchmark technology through the graph:
$T^t(x, y) = \{(x^t, y^t) : x^t \text{ can produce } y^t\} = \{(x^t, y^t) : y^t \le Y^t\lambda^t, x^t \ge x^t\lambda^t, z^t \in R_+^I\}$,
where Y^t and X^t are the observed output and input matrices in period t while λ^t is a vector of intensity variable. Following Shephard [9] the input distance function is defined at t as:
$D_i^t(x^t, y^t) = \sup\{\lambda : (x^t/\lambda, y^t) \in T^t\}$. Here we make use of the fact that the output distance function is reciprocal to the output-based Farrell measure of technical efficiency and compute for $\text{DMU}_o, o \in Q = \{1, ..., n\}$.

$$
\begin{array}{lll}
(D_i^t = (x_o^t, y_o^t))^{-1} = & \text{Min} \;\; \theta_o^t & \\
s.t. & \theta_o^t x_{io}^t \ge \sum_{j=1}^n \lambda_j^t x_{ij}^t, & i = 1, \ldots, m, \\
& y_{ro}^t \le \sum_{j=1}^n \lambda_j^t y_{rj}^t, & r = 1, \ldots, s, \\
& \lambda_j^t \ge 0, & j = 1, \ldots, n.
\end{array}
\tag{8}
$$

According to Chen et al. [3] and Zhu [11], we change model (8) into the following model:

$$
\begin{array}{lll}
(\widetilde{D}_i^t(x_o^t, y_o^t))^{-1} = & \text{Min} \;\; \theta_o^t & \\
s.t. & \sum_{j=1, j\neq o}^n \lambda_j^t \overline{x}_{ij}^t + \lambda_o^t \underline{x}_{io}^t \le \theta_o^t \underline{x}_{io}^t, & i \in BI, \\
& \sum_{j=1}^n \lambda_j^t x_{ij}^t \le \theta_o^t x_{io}^t, & i \notin BI \\
& \sum_{j=1, j\neq o}^n \lambda_j^t \underline{y}_{rj}^t + \lambda_o^t \overline{y}_{ro}^t \ge \overline{y}_{ro}^t, & r \in BO, \\
& \sum_{j=1}^n \lambda_j^t y_{rj}^t \ge y_{ro}^t, & r \notin BO \\
& \lambda_j^t \ge 0, & j = 1, \ldots, n.
\end{array}
\tag{9}
$$

In order to calculate the productivity of DMU_o between t and $t+1$, we need to solve four different linear programming problems for determining: $\widetilde{D}_i^t(x_o^t, y_o^t), \widetilde{D}_i^{t+1}(x_o^t, y_o^t), \widetilde{D}_i^t(x_o^{t+1}, y_o^{t+1})$ and $\widetilde{D}_i^{t+1}(x_o^{t+1}, y_o^{t+1})$. The computation of $\widetilde{D}_i^{t+1}(x_o^{t+1}, y_o^{t+1})$ is exactly like (9), where $t+1$ is substituted for t. This is provided in the following equivalent specification:

$$
\begin{array}{lll}
(\widetilde{D}_i^{t+1}(x_o^{t+1}, y_o^{t+1}))^{-1} = & \text{Min} \;\; \theta_o^{t+1} & \\
s.t. & \sum_{j=1, j\neq o}^n \lambda_j^{t+1} \overline{x}_{ij}^{t+1} + \lambda_o^{t+1} \underline{x}_{io}^{t+1} \le \theta_o^{t+1} \underline{x}_{io}^{t+1}, & i \in BI, \\
& \sum_{j=1}^n \lambda_j^{t+1} x_{ij}^{t+1} \le \theta_o^{t+1} x_{io}^{t+1}, & i \notin BI, \\
& \sum_{j=1, j\neq o}^n \lambda_j^{t+1} \underline{y}_{rj}^{t+1} + \lambda_o^{t+1} \overline{y}_{ro}^{t+1} \ge \overline{y}_{ro}^{t+1}, & r \in BO, \\
& \sum_{j=1}^n \lambda_j^{t+1} y_{rj}^{t+1} \ge y_{ro}^{t+1}, & r \notin BO, \\
& \lambda_j^{t+1} \ge 0, & j = 1, \ldots, n.
\end{array}
\tag{10}
$$

Two of the distance functions used to construct the Malmquist index requires information from two periods. The first of these is computed for DMU_o as:

$$\begin{array}{lll} (\widetilde{D}_i^t(x_o^{t+1}, y_o^{t+1}))^{-1} = & \text{Min} \;\; \theta_o^t & \\ s.t. & \sum_{j=1, j\neq o}^{n} \lambda_j^t \overline{x}_{ij}^t + \lambda_o^t \underline{x}_{io}^t \leq \theta_o^t \underline{x}_{io}^{t+1}, & i \in BI, \\ & \sum_{j=1}^{n} \lambda_j^t x_{ij}^t \leq \theta_o^t x_{io}^{t+1}, & i \notin BI, \\ & \sum_{j=1, j\neq o}^{n} \lambda_j^t \underline{y}_{rj}^t + \lambda_o^t \overline{y}_{ro}^t \geq \overline{y}_{ro}^{t+1}, & r \in BO, \\ & \sum_{j=1}^{n} \lambda_j^t y_{rj}^t \geq y_{ro}^{t+1}, & r \notin BO, \\ & \lambda_j^t \geq 0, & j = 1, \ldots, n. \end{array} \tag{11}$$

The last linear-programming problem we need to solve is also a mixed-period problem. It is specified as in (11), but the t and $t+1$ superscripts are transposed, as

$$\begin{array}{lll} (\widetilde{D}_i^{t+1}(x_o^t, y_o^t))^{-1} = & \text{Min} \;\; \theta_o^{t+1} & \\ s.t. & \sum_{j=1, j\neq o}^{n} \lambda_j^{t+1} \overline{x}_{ij}^{t+1} + \lambda_o^{t+1} \underline{x}_{io}^{t+1} \leq \theta_o^{t+1} \underline{x}_{io}^t, & i \in BI, \\ & \sum_{j=1}^{n} \lambda_j^{t+1} x_{ij}^{t+1} \leq \theta_o^{t+1} x_{io}^t, & i \notin BI, \\ & \sum_{j=1, j\neq o}^{n} \lambda_j^{t+1} \underline{y}_{rj}^{t+1} + \lambda_o^{t+1} \overline{y}_{ro}^{t+1} \geq \overline{y}_{ro}^t, & r \in BO, \\ & \sum_{j=1}^{n} \lambda_j^{t+1} y_{rj}^{t+1} \geq y_{ro}^t, & r \notin BO, \\ & \lambda_j^{t+1} \geq 0, & j = 1, \ldots, n. \end{array} \tag{12}$$

Following Fare et al. [5], we can calculate IDEA-based Malmquist productivity index as,

$$M_o = [\frac{\widetilde{D}_i^t(x_o^{t+1}, y_o^{t+1})}{\widetilde{D}_i^t(x_o^t, y_o^t)} \times \frac{\widetilde{D}_i^{t+1}(x_o^{t+1}, y_o^{t+1})}{\widetilde{D}_i^{t+1}(x_o^t, y_o^t)}]^{1/2}. \tag{13}$$

According to Zhu [11], we can change imprecise data of strong ordinal and ratio bounded into precise data and then calculate Malmquist index of such imprecise data.

4. Conclusion

The current paper calculates the Malmquist productivity index with data which some of them are imprecise. It is based upon the standard CCR model. It uses the standard linear DEA to perform IDEA-based productivity change analysis.

References

[1] Cave, D. W., L. R. Christensen and W. E. Diewert, Multilateral comparision of output, input, and productivity using superlative index numbers, Economic Journal, 92(365) (1982), 73-86.

[2] Charnes, A., W.W. Cooper and E. Rhodes, Measuring the efficiency of decision making units, European Journal of Operational Research, 2 (6) (1978), 429-444.

[3] Chen, Y., L. M. Seiford and J. Zhu, Imprecise data envelopment analysis, Management Science, (2002) (submitted).

[4] Cooper, W. W., K. S. Park and G. Yu, IDEA and AR-IDEA:Models for dealing With imprecise data in DEA. Management Science, 45 (1999), 597-607.

[5] Fare, R., S. Grosskopf, B. Lindgern and P. Roose, productivity change in Swedish Analysis, 3 (1992), 85-102. Pharmacies 1980-1989:A non-parametric Malmquist approach, Journal of productivity.

[6] Fare, R., S. Grosskopf, M. Norris and Z. Zhang, Productivity growth, technical progress, and efficiency changes in industrial country, American Economic Review ,84 (1994), 66-83.

[7] Kim, S. H. and K. S. Park, An application of data envelopment analysis in telephone offices evaluation with partial data, Computer Research, 26 (1999), 59-79.

[8] Malmquist, S., Index numbers and indifferent surfaces, Trabajos de Estadistica, 4(1) (1953), 209-42.

[9] Shephard, R. W., Theory of cost and production function, Princeton University Press, Princeton NJ, (1970).

[10] Zhu, J., Imprecise DEA via standard linear DEA models, Operations Research, (2002) (submitted).

[11] Zhu, J., Imprecise data envelopment analysis (IDEA): A review and improvement with an application, European Journal of Operational Research, 144 (2003), 513-529.

Received: May 27, 2007