A generalized super-efficiency model for ranking extreme efficient DMUs in stochastic DEA

M. Khoveyni[*], R. Eslami

Abstract In this current study a generalized super-efficiency model is first proposed for ranking extreme efficient decision making units (DMUs) in stochastic data envelopment analysis (DEA) and then, a deterministic (crisp) equivalent form of the stochastic generalized super-efficiency model is presented. It is shown that this deterministic model can be converted to a quadratic programming model. So far several approaches have been proposed in classic DEA by many researchers for ranking of efficient decision making units. In the previous proposed approaches, all inputs and outputs are respectively considered as deterministic (crisp) inputs and outputs while in real world, stochastic data may be present. It is necessary to mention that advantage of our proposed approach is capable of ranking of efficient decision making units with stochastic inputs and outputs. At last, an illustrative example highlights the proposed model and also a concluding comment, future extensions, and suggest possible future direction of research are all summarized.

1 Introduction

Ranking of efficient DMUs is an important question and many DEA researchers and practitioners have studied about it. Andersen and Petersen [1] were first addressed this question in their seminal paper where they introduced super-efficiency models to rank efficient decision making units. Also, Jahanshahloo et al. [2] introduced a generalized super-efficiency model for ranking of efficient DMUs in classic DEA. Stochastic DEA models have already been introduced in the literature; the research on generalized super-efficiency measure seems to have been solely focused on deterministic DEA models.

DEA efficiency measurement may be sensitive to stochastic variations in inputs and outputs. A decision making unit (DMU) which is measured as efficient relative to other DMUs may in turn be inefficient when such random variations are allowed. The need for a

---

[*] **Corresponding Author.**
    E-mail: mohammadkhoveyni@gmail.com (M. Khoveyni)

M. Khoveyni
    Department of Applied Mathematics, College of Basic Sciences, Yadegar-e-Imam Khomeini (RAH) Shahr-e-Rey Branch, Islamic Azad University, Tehran, Iran

R. Eslami
    Department of Mathematics, Faculty of Technology and Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran

flexible model which accounts for possible uncertainty in outputs and/or inputs and can produce robust results requires the introduction of a stochastic generalized super-efficiency model. Some DEA researchers have introduced stochastic formulation of the original DEA models for incorporating possible uncertainty in inputs and/or outputs [3-15]. In the previous proposed approaches, all inputs and outputs are respectively considered as deterministic (crisp) inputs and outputs while in real world, stochastic data may be present. It is necessary to mention that the robustness of efficiency results when input and output data are subject to stochastic measurement error was studied by Morita et al. [16], while a semi-infinite programming model in DEA was introduced by Jess et al. [17] for studying a chemical engineering problem.

Moreover, Asgharian et al. [18], Khodabakhshi [19], and Khodabakhshi and Asgharian [20] have studied stochastic input and output variations into DEA.

In this research, we first propose a generalized super-efficiency DEA model, allowing deterministic inputs and outputs to be stochastic. Then, a deterministic (crisp) equivalent form of our stochastic model is obtained and also, it is shown that this deterministic equivalent model can be transformed to a quadratic programming model. Note that advantage of our proposed approach is capable of ranking of efficient decision making units with stochastic inputs and outputs.

This paper is organized as follows: Generalized super-efficiency model in classic DEA is introduced in Section 2. In Section 3, the generalized super-efficiency model is presented in stochastic data envelopment analysis, and also, its deterministic (crisp) equivalent form is obtained. Furthermore, it is shown that the deterministic equivalent model of the stochastic generalized super-efficiency model can be converted to a quadratic program. A numerical example is presented in Section 4. Finally, in Section 5, the conclusion and some remarks are put forward.

2 Preliminaries

We consider n homogeneous DMUs $\left\{DMU_j \mid j = 1,...,n\right\}$ each having m inputs denoted by x_{ij} $(i = 1,...,m)$ and s outputs denoted by y_{rj} $(r = 1,...,s)$. Also, we assume that $\mathbf{X}_j = (x_{ij}) \in R^{m \times n}$ and $\mathbf{Y}_j = (y_{rj}) \in R^{s \times n}$ are non-negative deterministic elements. By using the variable return to scale (VRS), the production possibility set (PPS) is defined as follows:

$$T_c = \left\{ (\mathbf{X}, \mathbf{Y}) \left| \sum_{j=1}^{n} \lambda_j \mathbf{X}_j \leq \mathbf{X}, \sum_{j=1}^{n} \lambda_j \mathbf{Y}_j \geq \mathbf{Y}, \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, j = 1,2,...,n \right. \right\}. \tag{1}$$

Suppose that DMU_o $\left(o \in \{1,...,n\}\right)$ is one of the extreme efficient DMUs. By omitting DMU_o from T_c, we define the production possibility set T_c' as below:

$$T_c' = \left\{ (\mathbf{X}, \mathbf{Y}) \left| \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j \mathbf{X}_j \leq \mathbf{X}, \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j \mathbf{Y}_j \geq \mathbf{Y}, \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j = 1, \lambda_j \geq 0, j = 1,2,...,n, j \neq o \right. \right\}. \tag{2}$$

Generalized super-efficiency model is one of the basic models for ranking of efficient DMUs in classic DEA. This model was introduced by Jahanshahloo et al. [2]. It is used to rank DMU_o as follows:

$$Min \qquad \Gamma^o = \sum_{i=1}^{m} \alpha_i w_i + \sum_{r=1}^{s} \beta_r z_r$$

$$s.t. \qquad \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j x_{ij} \leq x_{io} + w_i, \quad i = 1,\ldots,m,$$

$$\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j y_{rj} \geq y_{ro} - z_r, \quad r = 1,\ldots,s,$$

$$\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j = 1, \tag{3}$$

$$\lambda_j \geq 0, \quad j = 1,\ldots,n, \ j \neq o,$$

$$w_i \geq 0, \quad i = 1,\ldots,m,$$

$$z_r \geq 0, \quad r = 1,\ldots,s,$$

where α_i $(i = 1,\ldots,m)$ and β_r $(r = 1,\ldots,s)$ are the positive weights that the manager gives us. Therefore model (3) is cooperated with the managers, because it considers the manager's opinions. It is necessary to mention that model (3) ranks only extreme efficient DMUs.

Note that in this study, "$*$" represents optimal solution values.

Theorem 1. *If $R^* = (\lambda^*, \mathbf{W}^*, \mathbf{Z}^*)$ be an optimal solution of the model (3), then all of the following constraints are active on R^* :*

$$(i) \ \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j x_{ij} \leq x_{io} + w_i, \quad i = 1,\ldots,m,$$

$$\tag{4}$$

$$(ii) \ \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j y_{rj} \geq y_{ro} - z_r, \quad r = 1,\ldots,s.$$

Proof. Refer to Jahanshahloo et al. [2]. □

Theorem 2. *Model (3) is always feasible and bounded.*

Proof. Refer to Jahanshahloo et al. [2]. □

Note that, the super-efficiency scores of the extreme efficient DMUs obtained by the above generalized super-efficiency model, model (3), can be ranked in a descending order. Obviously, the smallest of Γ^o s is corresponding with the last of extreme efficient DMU.

3 Stochastic data

Now, we are going to propose a generalized super-efficiency model in stochastic data envelopment analysis to rank extreme efficient DMUs.

3.1 Stochastic super-efficiency mode

In this section, we propose a stochastic model which allows for the possible presence of stochastic variability in the data. Following Cooper et al. [3], let $\tilde{\mathbf{X}}_j = (\tilde{x}_{1j},...,\tilde{x}_{mj})^t$, $\tilde{\mathbf{Y}}_j = (\tilde{y}_{1j},...,\tilde{y}_{sj})^t$ be random input and output related to DMU_j $(j=1,...,n)$. Let also $\mathbf{X}_j = (x_{1j},...,x_{mj})^t$ and $\mathbf{Y}_j = (y_{1j},...,y_{sj})^t$ show the corresponding vectors of expected values of inputs and outputs for DMU_j.

Note that in this research, the superscript "t" indicates a vector transpose.

Suppose that all input and output components are jointly normally distributed in the following chance constrained version of the stochastic model (3) with inequality constraints,

$$Min. \quad \Gamma^o = \sum_{i=1}^{m} \alpha_i w_i + \sum_{r=1}^{s} \beta_r z_r$$

$$s.t. \quad P\left(\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j \tilde{x}_{ij} \leq \tilde{x}_{io} + w_i \right) \geq 1-\alpha, \quad i=1,...,m,$$

$$P\left(\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j \tilde{y}_{rj} \geq \tilde{y}_{ro} - z_r \right) \geq 1-\alpha, \quad r=1,...,s,$$

$$\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j = 1, \tag{5}$$

$$\lambda_j \geq 0, \quad j=1,...,n, \ j \neq o,$$

$$w_i \geq 0, \quad i=1,...,m,$$

$$z_r \geq 0, \quad r=1,...,s.$$

where α is a predetermined value between 0 and 1 which specifies the significance level and P represents the probability measure.

The corresponding stochastic version of model (3), including slack variables is as follows:

$$Min \quad \Gamma^o = \sum_{i=1}^{m} \alpha_i w_i + \sum_{r=1}^{s} \beta_r z_r$$

$$s\,t. \quad P\left(\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j \tilde{x}_{ij} + s_i^- \leq \tilde{x}_{io} + w_i \right) = 1-\alpha, \quad i=1,\ldots,m,$$

$$P\left(\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j \tilde{y}_{rj} - s_r^+ \geq \tilde{y}_{ro} - z_r \right) = 1-\alpha, \quad r=1,\ldots,s,$$

$$\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j = 1,$$

$$\lambda_j \geq 0, \quad j=1,\ldots,n\,, j \neq o, \tag{6}$$

$$w_i \geq 0, \quad i=1,\ldots,m,$$

$$z_r \geq 0, \quad r=1,\ldots,s,$$

$$s_i^- \geq 0, \quad i=1,\ldots,m,$$

$$s_r^+ \geq 0, \quad r=1,\ldots,s.$$

In the next section, the deterministic equivalent form of the above stochastic super-efficiency model, model (5), is obtained.

3.2 Deterministic (crisp) equivalent model

In this section, we exploit the normality assumption to introduce a deterministic (crisp) equivalent form of the model (6). We need first recall a well-known fact about normally distributed random vectors that is used below. Suppose that $\vec{\mathbf{X}}_k \sim N\left(\vec{\boldsymbol{\mu}}_{k \times 1}, \boldsymbol{\Sigma}_{k \times k} \right)$, where $\vec{\boldsymbol{\mu}}_{k \times 1}$ and $\boldsymbol{\Sigma}_{k \times k}$ are respectively the mean value vector and the variance-covariance matrix. Then for any matrix $\mathbf{A}_{m \times k}$ we have $\mathbf{A}\vec{\mathbf{X}}_k \sim N\left(\mathbf{A}\vec{\boldsymbol{\mu}}, A\boldsymbol{\Sigma}A^t \right)$, where $\mathbf{A}^t$ is the transpose of $\mathbf{A}$. Using this results, one can obtain the following deterministic equivalent form of the stochastic generalized super-efficiency ranking model, model (6).

$$Min \quad \Gamma^o = \sum_{i=1}^{m} \alpha_i w_i + \sum_{r=1}^{s} \beta_r z_r$$

$$s.t. \quad \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j x_{ij} + s_i^- - \varphi^{-1}(\alpha)\sigma_i^I(\lambda) = x_{io} + w_i, \quad i = 1,\ldots,m,$$

$$\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j y_{rj} - s_r^+ + \varphi^{-1}(\alpha)\sigma_r^O(\lambda) = y_{ro} - z_r, \quad r = 1,\ldots,s,$$

$$\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j = 1,$$

$$\lambda_j \geq 0, \quad j = 1,\ldots,n, j \neq o, \tag{7}$$

$$w_i \geq 0, \quad i = 1,\ldots,m,$$

$$z_r \geq 0, \quad r = 1,\ldots,s,$$

$$s_i^- \geq 0, \quad i = 1,\ldots,m,$$

$$s_r^+ \geq 0, \quad r = 1,\ldots,s.$$

where φ is the cumulative distribution function (CDF) of a standard normal random variable and φ^{-1} is its inverse. We assume that x_{ij} and y_{rj} are the means of the input and output variables, which can be estimated by the observed values of the inputs and outputs. Using the aforementioned property of normal distribution, one can show that

$$\left(\sigma_i^I(\lambda)\right)^2 = \sum_{\substack{j=1 \\ j \neq o}}^{n} \sum_{\substack{k=1 \\ k \neq o}}^{n} \lambda_j \lambda_k Cov\left(\tilde{x}_{ij}, \tilde{x}_{ik}\right) + 2(\lambda_o - 1)\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j Cov\left(\tilde{x}_{ij}, \tilde{x}_{io}\right) + (\lambda_o - 1)^2 Var(\tilde{x}_{io}), \tag{8}$$

$$\left(\sigma_r^O(\lambda)\right)^2 = \sum_{\substack{k=1 \\ k \neq o}}^{n} \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_k \lambda_j Cov\left(\tilde{y}_{rk}, \tilde{y}_{rj}\right) + 2(\lambda_o - 1)\sum_{\substack{k=1 \\ k \neq o}}^{n} \lambda_k Cov\left(\tilde{y}_{rk}, \tilde{y}_{ro}\right) + (\lambda_o - 1)^2 Var(\tilde{y}_{ro}). \tag{9}$$

It is obvious, from the forms of $\sigma_i^I(\lambda)$ and $\sigma_r^O(\lambda)$, that model (7) is a non-linear program. We show that this non linear program can be transformed to a quadratic program. Suppose that q_i^I and q_r^O are non-negative variables. Replacing q_i^I and q_r^O, respectively, $\sigma_i^I(\lambda)$ and $\sigma_r^O(\lambda)$ then adding the following quadratic equality constraints

$$\left(q_i^I\right)^2 = \left(\sigma_i^I(\lambda)\right)^2, \tag{10}$$

$$\left(q_r^O\right)^2 = \left(\sigma_r^O(\lambda)\right)^2. \tag{11}$$

Model (7) is transformed to a quadratic programming problem. Therefore, we obtain the optimal values w_i^*, z_r^*, s_i^{-*}, and s_r^{+*} by solving the quadratic program. Finally, we have the following deterministic equivalent form of the stochastic generalized super-efficiency ranking model.

$$Min \quad \Gamma^o = \sum_{i=1}^{m} \alpha_i w_i + \sum_{r=1}^{s} \beta_r z_r$$

$$s.t. \quad \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j x_{ij} + s_i^- - \varphi^{-1}(\alpha) q_i^I = x_{io} + w_i, \quad i = 1, \ldots, m,$$

$$\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j y_{rj} - s_r^+ + \varphi^{-1}(\alpha) q_r^O = y_{ro} - z_r, \quad r = 1, \ldots, s,$$

$$\left(q_i^I\right)^2 = \sum_{\substack{j=1 \\ j \neq o}}^{n} \sum_{\substack{k=1 \\ k \neq o}}^{n} \lambda_j \lambda_k Cov(\tilde{x}_{ij}, \tilde{x}_{ik}) + 2(\lambda_o - 1) \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j Cov(\tilde{x}_{ij}, \tilde{x}_{io}) + (\lambda_o - 1)^2 Var(\tilde{x}_{io}),$$

$$\left(q_r^O\right)^2 = \sum_{\substack{k=1 \\ k \neq o}}^{n} \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_k \lambda_j Cov(\tilde{y}_{rk}, \tilde{y}_{rj}) + 2(\lambda_o - 1) \sum_{\substack{k=1 \\ k \neq o}}^{n} \lambda_k Cov(\tilde{y}_{rk}, \tilde{y}_{ro}) + (\lambda_o - 1)^2 Var(\tilde{y}_{ro}),$$

$$\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j = 1, \tag{12}$$

$$\lambda_j \geq 0, \quad j = 1, \ldots, n, \ j \neq o,$$

$$w_i \geq 0, \quad i = 1, \ldots, m,$$

$$z_r \geq 0, \quad r = 1, \ldots, s,$$

$$s_i^- \geq 0, \quad i = 1, \ldots, m,$$

$$s_r^+ \geq 0, \quad r = 1, \ldots, s.$$

4 Numerical example

In this example, we are going to rank extreme efficient decision making units with one input and one output by proposed model. The average values of the input and output are used over the two years for each of the 4 DMUs as the expected input and output of each DMU (see Table 1). With using CCR model, we realize that DMU_1, DMU_2, and DMU_3 are extreme efficient DMUs.

Table 1 Average of data for two years.

DMU	Input	Output
1	2	1
2	3	4
3	4	5
4	3	2

Assume that $\alpha \approx 0.4$ then $\varphi^{-1}(\alpha) = -0.2124$ and also, we consider $\alpha_1 = 1$ and $\beta_1 = 1$. For simplicity and computational convenience, it is assumed that all DMUs have the same variances in the input and output. The variances for the output and the input can therefore be estimated by $Var(\tilde{y}) = \frac{1}{3} \sum_{j=1}^{4} (y_j - \bar{y})^2 = \frac{10}{3}$ and $Var(\tilde{x}) = \frac{1}{3} \sum_{j=1}^{4} (x_j - \bar{x})^2 = \frac{2}{3}$ where $\bar{y} = \frac{1}{4} \sum_{j=1}^{4} y_j = 3$ and $\bar{x} = \frac{1}{4} \sum_{j=1}^{4} x_j = 3$. Note that, y_j and x_j are the average outputs and

inputs of each DMU_j between two years as presented in Table 1. Also, it is assumed that outputs and inputs of different DMUs are independent each other. This independent assumption then implies that $Cov(\tilde{y}_k, \tilde{y}_j) = 0$ and $Cov(\tilde{x}_j, \tilde{x}_k) = 0$. Therefore, model (12) can be converted to linear program which can be solved using the simplex method. The stochastic results obtained from GAMS software are presented in Table 2.

Table 2 Computational results for stochastic model, model (12).

DMU	λ_1^*	λ_2^*	λ_3^*	λ_4^*	s^{-*}	s^{+*}	w^*	z^*	Γ^*
1	----	1.000	0	0	0	2.612	1.173	0	1.173
2	0.153	----	0.847	0	0	0	0.867	0	0.867
3	0	1.000	----	0	0.827	0	0	1.388	1.388

In this study, a generalized super-efficiency model is presented for ranking extreme efficient DMUs in stochastic data envelopment analysis based upon the largeness of Γ^* value in Table 2. According to Table 2, all of the 3 CCR extreme efficient DMUs are classified by using our presented model, model (12), that DMU_3 is as the best extreme efficient DMU.

5 Conclusions

Stochastic models may be better suited for DEA when there is uncertainty associated with the inputs and/or outputs of DMUs or when an analyst may be wondering how much change can be incurred in the ranking of DMUs if some of the inputs and/or outputs change. In this research, we have proposed a generalized super-efficiency model for ranking extreme efficient DMUs in stochastic data envelopment analysis. Also, we have developed a stochastic version of the proposed generalized super-efficiency model and obtained a deterministic (crisp) equivalent form of the stochastic version. This deterministic equivalent model can be converted to a quadratic problem. Note that, we can use our proposed model with constant return to scale. Appling the proposed approach in different fields, practically, would be interesting for further research.

References

1. Andersen, P., Petersen, N. C., (1993). A procedure for ranking efficient units in data envelopment analysis. Management Science, 39 (10), 1261-1264.
2. Jahanshahloo, G. R., Hosseinzadeh Lotfi, F., Mehrabian, S., Fallah Jelodar, M., (2007). Generalized super-efficiency model for ranking efficient DMUs in data envelopment analysis. (Working paper).
3. Cooper, W. W., Deng, H., Huang, Z., Li, S. X., (2004). Chance constrained programming approaches to congestion in stochastic data envelopment analysis. European Journal of Operational Research, 155 (2), 487-501.
4. Cooper, W. W., Seiford, L. M., Tone, K., (2007). Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software (Second Edition). New York, Springer Science+Business Media: Publisher.
5. Huang, Z., Li, S. X., (1996). Dominance stochastic models in data envelopment analysis. European Journal of Operation Research, 95 (2), 390-403.

6. Land, K. C., Lovell, C. A. K., Thore, S., (1988). Chance constrained data envelopment analysis. Managerial and Decision Economics, 14, 541-554.
7. Olesen, O. B., Petersen, N. C., (1995). Chance constrained efficiency evaluation. Management Science, 41, 442-457.
8. Khodabakhshi, M., Aryavash, K., (2012). Ranking all units in data envelopment analysis. Applied Mathematics Letters, 25 (12), 2066-2070.
9. Zerafat Angiz, M., Mustafa, A., Kamali, M. J., (2013). Cross-ranking of Decision Making Units in Data Envelopment Analysis. Applied Mathematical Modelling, 37 (1–2), 398-405.
10. Guo, D., Wu, J., (2013). A complete ranking of DMUs with undesirable outputs using restrictions in DEA models. Mathematical and Computer Modelling, 58 (5–6), 1102-1109.
11. Eslami, R., Khoveyni, M., (2011). A New Method for Ranking Extreme Efficient DMUs Based on Changing the Reference Set Using L2 – Norm. International Journal of Applied Operational Research, 1 (1), 65-71.
12. Zamanian, G. R., Shahabinejad, V., Yaghoubi, M., (2013). Application of DEA and SFA on the Measurement of Agricultural Technical efficiency in MENA1 Countries. International Journal of Applied Operational Research, 3 (2), 43–51.
13. Kordrostami, S., Amirteimoori, A., Pakdaman Ashrafiyeh, M., (2012). Slack-Based measurement with Rough Data. International Journal of Applied Operational Research, 2 (2), 59–65.
14. Rahimi, I., Behmanesh, R., (2012). Improve poultry farm efficiency in Iran: using combination neural networks, decision trees, and data envelopment analysis (DEA). International Journal of Applied Operational Research, 2 (3), 69–84.
15. Azizi, H., Jahed, R., (2011). An Improvement for efficiency Interval: Efficient and Inefficient Frontiers. International Journal of Applied Operational Research, 1 (1), 49–63.
16. Hiroshi, M., Seiford, L. M., (1999). Characteristics on stochastic DEA efficiency. Journal of the Operational Research Society, 42 (4), 389-404.
17. Jess, A., Jongen, H. Th., Neralic, L., Stein, O., (2001). Semi-infinite programming model in data envelopment analysis. Optimization, 49 (4), 369-385.
18. Asgharian, M., khodabakhshi, M., Neralic, L., (2010). Congestion in stochastic data envelopment analysis: An input relaxation approach. International Journal of Statistics and Management System, 5 (1-2), 84-106.
19. Khodabakhshi, M., (2007). A super-efficiency model based on improved outputs in data envelopment analysis. Applied Mathematics and computation, 184 (2), 695-703.
20. Khodabakhshi, M., Asgharian, M., (2009). An input relaxation measure of efficiency in stochastic data envelopment analysis. Applied Mathematical Modelling, 33, 2010-2023.