No. **636**

A note on DEA computation: The use of the exclusion model

by

Akihiro Hashimoto

July 1995

A note on DEA computation: The use of the exclusion model

Akihiro Hashimoto

Institute of Socio-Economic Planning, University of Tsukuba, Tsukuba, Ibaraki 305, Japan

Abstract

In the DEA (Data Envelopment Analysis) computation with n DMUs (Decision Making Units), we would usually like to obtain not only DEA efficiency scores of DMUs but also some other information, e.g., whether each DMU is on the efficient frontier, on the extended frontier or not, etc. To get such information, we cannot make do with solving n LP (Linear Programming) problems but must also solve considerably many additional LP problems in using the usual LP softwares. This paper shows that we can get the DEA information through solving nearly n problems by using the DEA exclusion model instead of the standard DEA model. This is a merit other than discriminating DEA efficient DMUs to use the DEA exclusion model.

Keywords: Data envelopment analysis; DEA computation; DEA exclusion model

1. DEA computation

DEA (Data Envelopment Analysis) measures the relative efficiency of DMUs (Decision Making Units). The model to obtain DEA efficiency score h_{j0}^*, $0 < h_{j0}^* \leq 1$, for target DMU j_0 is expressed as follows, converted into the LP (Linear Programming) formulation (Charnes *et al.* [3]):

$$\text{Maximize} \quad h_{j_0} = \sum_{r=1}^{t} u_r y_{rj_0} \tag{1.1a}$$

$$\text{subject to} \quad \sum_{i=1}^{m} v_i x_{ij_0} = 1, \tag{1.1b}$$

$$\sum_{r=1}^{t} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \quad j = 1, \ldots, n, \tag{1.1c}$$

$$u_r, \; v_i \geq 0, \quad r = 1, \ldots, t, \; i = 1, \ldots, m, \tag{1.1d}$$

where y_{rj} = the amount of output r from DMU j; x_{ij} = the amount of input i to DMU j; u_r = the weight given to output r; v_i = the weight given to input i; n = the number of DMUs; t = the number of outputs; m = the number of inputs; h_{j0}^* = the maximum of h_{j0}. Solving this problem for each DMU j_0, $j_0 = 1, \ldots, n$, we get DEA scores for all the DMUs.

The dual of problem (1.1) is as follows:

$$\text{Minimize} \quad \theta \tag{1.2a}$$

$$\text{subject to} \quad \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ = y_{rj_0}, \quad r = 1, \ldots, t, \tag{1.2b}$$

$$x_{ij_0}\theta - \sum_{j=1}^{n} x_{ij}\lambda_j - s_i^- = 0, \quad i = 1, \ldots, m, \tag{1.2c}$$

$$\lambda_j, \; s_r^+, \; s_i^- \geq 0, \quad j = 1, \ldots, n, \; r = 1, \ldots, t, \; i = 1, \ldots, m, \tag{1.2d}$$

$$(\theta \text{ unconstrained}),$$

where θ, λ_j = dual variables; s_r^+, s_i^- = slack variables. Of course, the minimum $\theta^* = h_{j0}^*$ at an optimum.

The set of all DMUs is partitioned into the following four subsets: E, E', F and N (Charnes *et al.* [4, 5], Sueyoshi [10]). Here, E = a set of DEA efficient DMUs with DEA score $\theta^* = 1$, each of which is a vertex on the efficient frontier; E' = a set of DEA efficient DMUs with $\theta^* = 1$, each of which is on the efficient frontier but not a vertex; F = a set of DEA inefficient DMUs with $\theta^* = 1$, each of which is on the extended frontier; N = a set of DEA inefficient DMUs with $\theta^* < 1$, each of which is not on the efficient nor extended frontier. Denoting n_E = the number of DMUs belonging to subset E, etc., $n = n_E + n_{E'} + n_F + n_N$. We should here note that in the general DEA analysis, $n_{E'}$ or n_F is considerably small as compared with n_E or n_N. Fig. 1 shows the four subsets of DMUs in the case of two inputs and one output for graphical simplicity.

Through DEA computation, we would like to obtain the following information: (1) DEA score and the optimal weights u_r^*, v_i^* for each DMU; (2) which subset each DMU belongs to; (3) for each of the DEA inefficient DMUs belonging to subset F or N, reference set, i.e., the set of DMUs being efficient with the weights optimal to the inefficient target DMU, and combination coefficients. Usually, we get such DEA information by going through the following steps using problem (1.2) [4, 5]:

Step 1.1. Solving problem (1.2) for each DMU j_0 of the n DMUs, we obtain DEA score θ^* and the optimal weights u_r^*, v_i^* as shadow prices of constraints (1.2b) and (1.2c). We can here identify subset N in terms of $\theta^* < 1$. However, for DMU $j_0 \in N$, DMUs j with $\lambda_j^* > 0$ at this step might not necessarily be elements of the reference set because they might be DEA inefficient DMUs on the extended frontier (see Step 1.3).

Step 1.2. For each DMU j_0 of the $(n_E + n_{E'} + n_F)$ DMUs with $\theta^* = 1$ at Step 1.1, we solve a modification of problem (1.2) in which the objective "Minimize θ" is replaced by

$$\text{Maximize} \quad \sigma_{j_0} = \sum_{r=1}^{t} s_r^+ + \sum_{i=1}^{m} s_i^- \tag{1.3}$$

and θ is fixed at $\theta = 1$. We can here identify subset F in terms of the maximum $\sigma_{j0}^* > 0$ and obtain the reference set and combination coefficients for DMUs $j_0 \in F$ in terms of $\lambda_j^* > 0$ of this step.

Step 1.3. Suppose that the DMUs belonging to subset F identified at Step 1.2 are included in the set of DMUs j with $\lambda_j^* > 0$ for DMUs $j_0 \in N$. Then, for each of such DMUs $j_0 \in N$, we again solve a modified form of problem (1.2) in which all λ_j for $j \in F$ are fixed at $\lambda_j = 0$ and obtain the reference set and combination coefficients in terms of $\lambda_j^* > 0$ of this step. Denoting n_α = the number of problems to be solved here, n_α would be very small. For each of the remaining $(n_N - n_\alpha)$ DMUs $j_0 \in N$, the DMUs j with $\lambda_j^* > 0$ at Step 1.1 form the reference set.

Step 1.4. For each DMU j_0 of the $(n_E + n_{E'})$ DMUs with $\sigma_{j0}^* = 0$ at Step 1.2, we solve a modified form of problem (1.2), this time with the objective "Minimize λ_{j0}" and again with θ fixed at $\theta = 1$. If the minimum $\lambda_{j0}^* = 0$, then the DMU j_0 belongs to subset E', otherwise it belongs to E.

In this way, to obtain the DEA information, we must solve considerably many LP problems in addition to the n LP problems to be solved at Step 1.1, i.e., a total of $(n + 2n_E + 2n_{E'} + n_F + n_\alpha)$ LP problems to be solved throughout Steps 1.1–1.4. Especially at Step 1.2, we solve the problem for each of the $(n_E + n_{E'} + n_F)$ DMUs with $\theta^* = 1$ to identify subset F. In the general DEA analysis, though fairly many DMUs have the value $\theta^* = 1$, n_F is much smaller than n_E as mentioned before. Therefore, we do computation at Step 1.2 while expecting that elements of the subset F would scarcely be found. In fact, in a DEA case with 47 DMUs, four inputs and four outputs (Hashimoto and Ishikawa [7]), we had to solve 26 problems at Step 1.2 in addition to 47 problems at Step 1.1, and eventually found no DMUs belonging to F. Further, if we would like to discriminate between subsets E and E', we must solve 26 more problems at Step 1.4, i.e., a total of 99 problems throughout Steps 1.1–1.4.

Of course, we can get such all DEA information by going through only Step 1.1, i.e., through only n LP problems, if we solve for each of the n DMUs a modification of problem (1.2) in which the objective function is replaced by

$$\text{Minimize} \quad \theta - \varepsilon\left(\sum_{r=1}^{t} s_r^+ + \sum_{i=1}^{m} s_i^-\right) + \delta\lambda_{j_0}, \tag{1.4}$$

where ε, δ = positive non-Archimedean infinitesimals, $1 \gg \varepsilon \gg \delta$, i.e., ε is preemptively greater than δ. This can be done in terms of LP computer programs coded taking three preemptive priority factors into consideration like the two-phase simplex method using artificial variables (see, e.g., Lee [8]). However, such programs would cost much labor to be developed and generally be unavailable. This paper considers a way to reduce the load of DEA computation when we do it in terms of the usual LP softwares, e.g., LINDO (Schrage [9]).

2. Using the DEA exclusion model

We use the DEA exclusion model for reducing DEA computation load. The DEA exclusion model was proposed by Andersen and Petersen [2] as one that discriminates DEA efficient DMUs. In this model, the DMU being evaluated is excluded from the comparison set (Adolphson *et al.* [1]). (See Hashimoto [6] for an application.)

The DEA exclusion model corresponding to model (1.2) is expressed as follows:

$$\text{Minimize} \quad \phi \tag{2.1a}$$

$$\text{subject to} \quad \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ = y_{rj_0}, \quad r = 1, \ldots, t, \tag{2.1b}$$

$$x_{ij_0}\phi - \sum_{j=1}^{n} x_{ij}\lambda_j - s_i^- = 0, \quad i = 1, \ldots, m, \tag{2.1c}$$

$$\lambda_{j_0} = 0, \tag{2.1d}$$

$$\lambda_j,\ s_r^+,\ s_i^- \geq 0, \quad j = 1, \ldots, n,\ r = 1, \ldots, t,\ i = 1, \ldots, m, \tag{2.1e}$$

(ϕ unconstrained),

where ϕ = variable used instead of θ in model (1.2). This is different from (1.2) by only constraint (2.1d). Since the region of feasible solutions of the exclusion model is more strictly constrained than that of the standard model (1.2) by (2.1d), the minimum $\phi^* \geq \theta^*$.

The DEA score θ^* for DMU j_0 is the ratio of vector length of DMU j_0's reference point on the frontier to that of DMU j_0 itself. Since DMU j_0 belonging to subset E, E' or F is on the frontier, the DMU itself and its reference point are identical. Therefore, $\theta^* = 1$ for DMU $j_0 \in E$, E' or F and $\theta^* < 1$ for DMU $j_0 \in N$ as mentioned in Sec. 1. (See Fig. 1.)

On the other hand, the DEA exclusion score ϕ^* has the similar implication to the DEA score θ^*. However, excluding the DMU j_0 being evaluated from the comparison set in the exclusion model, the frontier and the reference point on it shift for DMU $j_0 \in E$. For DMU $j_0 \in E'$, F or N, the frontier does not shift, so that the reference point does not change. Therefore, $\phi^* > \theta^* = 1$ for DMU $j_0 \in E$; $\phi^* = \theta^* = 1$ for DMU $j_0 \in E'$ or F; $\phi^* = \theta^* < 1$ for

DMU $j_0 \in N$ [2].

Using the DEA exclusion model (2.1), we can get the DEA information by going through the following steps:

Step 2.1. Solving problem (2.1) for each DMU j_0 of the n DMUs, we obtain DEA exclusion score ϕ^* and shadow prices. We can here identify subset E in terms of $\phi^* > 1$ and subset N in terms of $\phi^* < 1$.

Step 2.2. For each DMU j_0 of the $(n_{E'} + n_F)$ DMUs with $\phi^* = 1$ at Step 2.1, we solve a modification of problem (2.1) in which the objective "Minimize ϕ" is replaced by (1.3) and ϕ is fixed at $\phi = 1$. We can here identify subset E' in terms of the maximum $\sigma_{j0}^* = 0$ and subset F in terms of $\sigma_{j0}^* > 0$, and obtain the reference set and combination coefficients for DMUs $j_0 \in F$.

Step 2.3. Like Step 1.3, for each DMU $j_0 \in N$ with $\lambda_j^* > 0$ for $j \in F$, we solve a modified form of problem (2.1) in which all λ_j for $j \in F$ are fixed at $\lambda_j = 0$. Through this step and Step 2.1, we can get the reference set and combination coefficients for DMUs $j_0 \in N$.

We should here note the following: (1) The DEA score θ^* for DMU $j_0 \in E$ with $\phi^* > 1$ is $\theta^* = 1$. Of course, $\theta^* = \phi^*$ for others. (2) The optimal weights u_r^*, v_i^* corresponding to $\theta^* = 1$ for DMU $j_0 \in E$ are calculated as u_r^* = [shadow price of constraint (2.1b)] / ϕ^*; v_i^* = shadow price of constraint (2.1c). Because the dual of problem (2.1) is a modification of problem (1.1) in which constraint (1.1c) for $j = j_0$ is excluded. The shadow prices obtained at Step 2.1 are the optimal weights as they are for DMU $j_0 \in E'$, F or N. (3) The step corresponding to Step 1.4 is not needed.

Since the number of problems to be solved at Step 2.3 may be considered equal to that at Step 1.3, we would solve $(n + n_{E'} + n_F + n_\alpha)$ problems throughout Steps 2.1–2.3. It means $(2n_E + n_{E'})$ problems reduction from Steps 1.1–1.4, and $2n_E$ is fairly many. Further, $(n_{E'} + n_F + n_\alpha)$ would be very small, so that we may solve nearly n problems to get the DEA information by using the exclusion model. In the DEA case mentioned in Sec. 1, since $n_{E'} = n_F = 0$ (i.e., $n_\alpha = 0$) in fact, we could reduce the number of problems to be solved from 99 of Steps 1.1–1.4 to 47 (= n) of Steps 2.1–2.3.

Of course, at Steps 2.1–2.2, we also obtain the information of discriminating DEA efficient DMUs [2]. Therefore, we propose to use the DEA exclusion model for DEA computation as one getting more DEA information through less DEA computation than Steps 1.1–1.4.

References

[1] D. L. Adolphson, G. C. Cornia and L. C. Walters, "A unified framework for classifying DEA models", *Operational Research '90* (H. E. Bradley, ed.), Pergamon Press, Oxford, 647-657 (1991).

[2] P. Andersen and N. C. Petersen, "A procedure for ranking efficient units in data envelopment analysis", *Mgmt Sci.* **39**, 1261–1264 (1993).

[3] A. Charnes, W. W. Cooper and E. Rhodes, "Measuring the efficiency of decision making units", *Eur. J. Opl Res.* **2**, 429–444 (1978).

[4] A. Charnes, W. W. Cooper and R. M. Thrall, "Classifying and characterizing efficiencies and inefficiencies in data envelopment analysis", *Opns Res. Lett.* **5**, 105–110 (1986).

[5] A. Charnes, W. W. Cooper and R. M. Thrall, "A structure for classifying and character–izing efficiency and inefficiency in data envelopment analysis", *J. Productiv. Anal.* **2**,

197–237 (1991).
[6] A. Hashimoto, "A ranked voting system using DEA/AR exclusion model: A note", Inst. Socio-Econ. Plann. Discussion Pap. Ser. 626, Univ. Tsukuba, Tsukuba, 1995.
[7] A. Hashimoto and H. Ishikawa, "Using DEA to evaluate the state of society as measured by multiple social indicators", *Socio-Econ. Plann. Sci.* **27**, 257–268 (1993).
[8] S. M. Lee, *Goal Programming for Decision Analysis*, Mason & Lipscomb Publishers, New York, 1972.
[9] L. Schrage, *LINDO: An Optimization Modeling System* (4th ed.), Scientific Press, South San Francisco, 1991.
[10] T. Sueyoshi, "Measuring the industrial performance of Chinese cities by data envelopment analysis", *Socio-Econ. Plann. Sci.* **26**, 75–88 (1992).

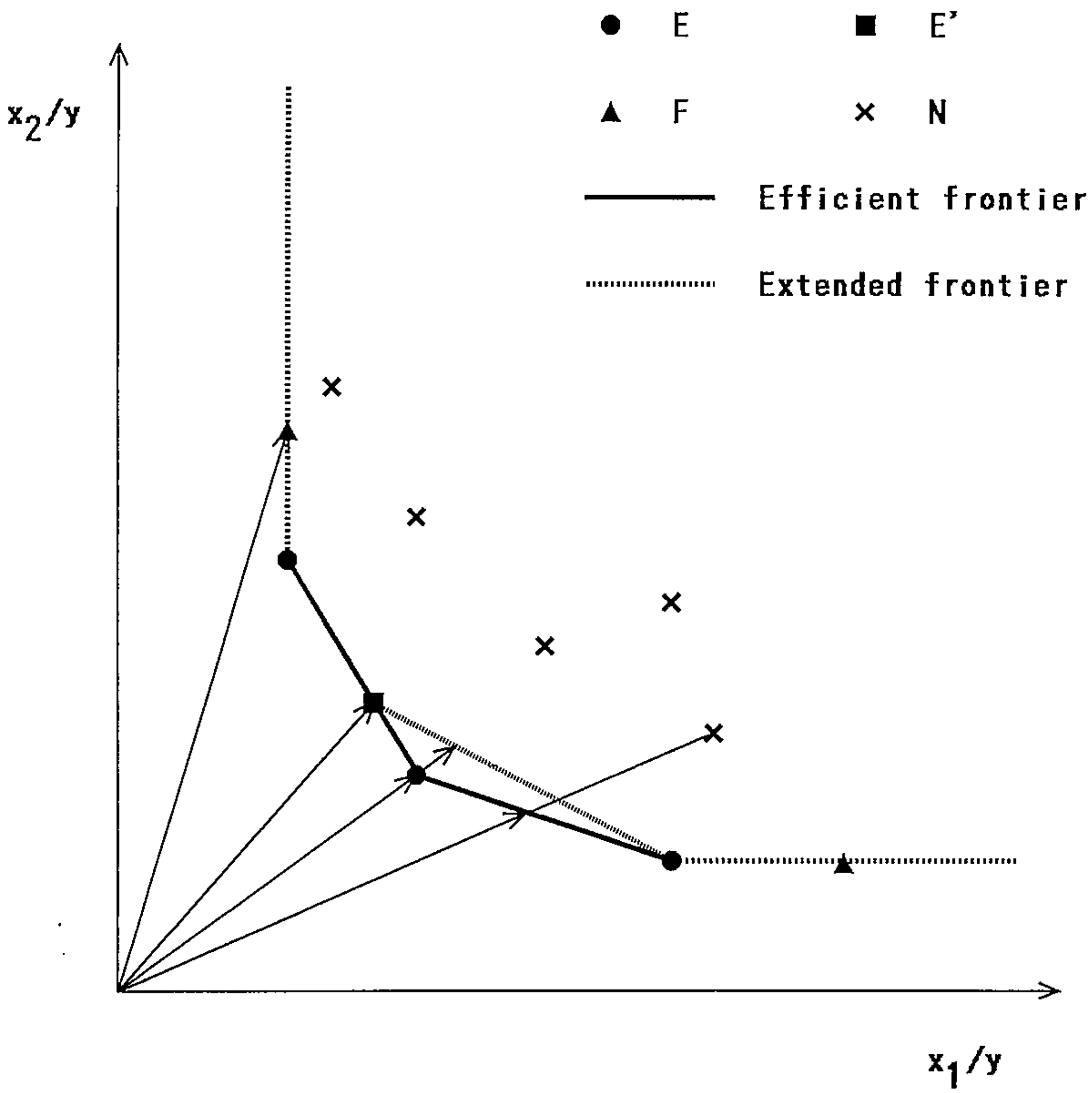

Fig. 1. DMUs belonging to the four subsets, reference point vectors on the frontier and their shifts.