The Comparison of Stochastic and Deterministic DEA Models

Michal Houda, Jana Klicnarová

Abstract: The idea of deterministic data envelopment analysis (DEA) models is well-known. If we start to study the stochastic DEA models, it seems at first glance, that the ideas of these two models are totally different. The aim of this paper is to explain that there is a really nice connection between the ideas of stochastic and deterministic models.

Key words: Data Envelopment Analysis · Deterministic DEA · Stochastic DEA · Efficiency

JEL Classification: C44 · C61

1 Introduction

The idea of the piecewise-linear convex hull as a frontier estimation for effective units goes back to Farrell (1957) and it was consider by a few authors in following years. The main attention received this topic when Charnes, Cooper and Rhodes (1978) presented their paper, where the Data Envelopment Analysis (DEA) was introduced. Since then there were published a large number of papers and books which extended these results and DEA methodology – see for example Coelli (2009).

The basic idea of DEA is the comparison of some units which are characterized by several inputs and outputs. The main aim of this analysis is to identify so-called efficient units. More precisely, the aim is to estimate frontier function and to measure the efficiencies of units relative to this estimated frontier. The other possible point of view at this problem comes from the decision making theory. In fact, we have some units and we know there evaluation according to several criteria, some of them are cost type, some of them are benefit type. Our aim is to find such units (alternatives in the terms of multiple attribute decision making), which are not dominated and which are Pareto efficient. The solving of DEA problems leads to the problem of mathematical programming, in case of deterministic models to linear programming, in case of stochastic model to non-linear programming problems.

The paper is organized as follows. In the section 2 we introduce the basic notation necessary for the setting of these problems. Section 3 introduces the deterministic DEA models and the stochastic DEA models are given in section 4. To show the connection between deterministic and stochastic models, we present the deterministic models in non-classical way. We do not use a formulation based on the idea of optimization of efficiency – which is widely used for deterministic models, but we use a dual approach – an approach of dominated alternatives. This allows us to show a close connection between deterministic and stochastic formulation of the problem.

2 Methods

To set up the models, let us introduce the following notation. By $DMU_k$, we denote a $k$-th decision making unit, where $k = 1, ..., K$ and $K$ is a number of units in question. Then, we write $X := (x_{ik}) \in \mathbb{R}^{m \times K}$ for input matrix where each column $x_k = (x_{1k}, ..., x_{mk})^T$ represents the input vector of $k$-th decision unit. On the other hand, the rows of the matrix $X$ represent the values of individual inputs, that is, $x_i = (x_{i1}, ..., x_{ik})^T$ gives the values for the $i$-th input of all units. In the same way, we use $Y := (y_{jk}) \in \mathbb{R}^{n \times K}$ as the output matrix, and $y_k, y_j$ the corresponding output vectors (column and row vectors).

The key object in our investigation represents the production possibility set, denoted $PPS$ in short. It is the set of all possible/allowed combination of inputs and outputs. This set is given by the data, it varies from analysis to analysis. It is also important to remark that – as mentioned above – matrices $X$ and $Y$ are not variables (as usual in optimization models) but they represent the data (i.e., inputs and outputs of analyzed units).

The analysis of the efficiency of a unit is closely related to the definition of the dominance property. As we will see, its definition varies depending upon the production possibility set chosen. As a consequence, the efficiency of the unit

Mgr. Michal Houda, Ph.D, University of South Bohemia in České Budějovice, Faculty of Economics, Department of Applied Mathematics and Informatics, Studentská 13, CZ-37005 České Budějovice, e-mail: houda@ef.jcu.cz

RNDr. Jana Klicnarová, Ph.D, University of South Bohemia in České Budějovice, Faculty of Economics, Department of Applied Mathematics and Informatics, Studentská 13, CZ-37005 České Budějovice, e-mail: janaklic@ef.jcu.cz
in question will be defined individually for each model considered and cannot be seen as a “universal” property of the unit.

The unit which is considered for efficiency, will be denoted $DMU_0$ through the paper.

3 Discrete and Continuous Production Possibility Sets

In this section, we describe traditional deterministic DEA models appearing in the literature. Apart from the usual economical formulations (based on comparing the input/output ratio), we concentrate on the characterization of these models based on production possibility sets, which later facilitates the transition to their stochastic versions.

3.1 Discrete Production Possibility Set – Additive Model

We start our investigation with the production possibility set composed from only the actually observed units, that is

$$PPS_i := \{(x_1, y_1), \ldots, (x_K, y_K)\}.$$  \hspace{1cm} (1)

To characterize the efficiency of the units with respect to $PPS_i$, we perform a simple pairwise comparison between units using the following definition of efficiency dominance.

**Definition 1** (Cooper, Huang, Li, 2004) We say that $DMU_1$ dominates $DMU_2$ with respect to $PPS_i$ if $x_1 \leq x_2$ and $y_1 \geq y_2$ with strict inequality holding for at least one of the components in the input or output vector. We say that $DMU_0$ is efficient with respect to $PPS_i$ if there is no unit dominating $DMU_0$ with respect to $PPS_i$.

Bowlin et al. (1984) formulated a so-called additive model with binary constraints as follows:

$$\begin{align*}
\text{maximize} & \quad (\sum_i s_i^- + \sum_j s_j^+) \\
\text{subject to} & \quad \sum_k x_{ik} \lambda_k + s_i^- = x_{i0} \text{ for all inputs } i = 1, \ldots, m, \\
& \quad \sum_k y_{jk} \lambda_k - s_j^+ = y_{j0} \text{ for all outputs } j = 1, \ldots, n, \\
& \quad \sum_k \lambda_k = 1, \lambda_k \in \{0, 1\}^K, s_i^-, s_j^+ \geq 0. \hspace{1cm} (2)
\end{align*}$$

The decision variables $s_i^-$, $s_j^+$ are the slack and surplus variables, respectively, for the inequalities $X\lambda \leq x_0$, $Y\lambda \geq y_0$. (We will generally refer to them as slacks in the rest of the paper.) As there can be only one element of the binary variables equal to one (say $k$-th, that is $\lambda_k = 1$), this element identifies the unit (namely $DMU_k$) for pairwise comparison with $DMU_0$, at each solution (iteration). Maximizing slacks ensures that, at an optimal solution, $DMU_k$ is the “most dominant” unit, with respect to $DMU_0$. At optimum, if all slacks are zero, then $DMU_k$ does not dominate $DMU_0$ in sense of Definition 1, hence there is no unit dominating $DMU_0$, henceforth $DMU_0$ is efficient. Figure 1 illustrates the procedure: the most dominating unit for $F$ is the unit $B$. As the slacks and surpluses are nonzero in this case, $F$ is not efficient. The units $A$, $B$, $E$, and $H$ do not have dominating units in sense of Definition 1, hence they are efficient.

**Figure 1** Example of the additive DEA model

![Figure 1](source: Own processing of the data from Table 1.1 in Cooper, Seiford, Tone (2007))
Due to the presence of integrality constraints, this optimization problem unfortunately does not have its dual form, that is, with objective in a traditional input/output ratio form. Its continuous relaxation leads to the model which is described in the following section.

3.2 Continuous (Convex) Production Possibility Set – BCC Model

A continuous relaxation of (2), given originally by Banker, Cooper, Charnes (1984) and known as BCC model, supposes the production possibility set to be the convex hull of observed units. That is,

\[ PPS_C = \{(x, y)|x = X\lambda, y = Y\lambda, \lambda = (\lambda_1, \lambda_2, ..., \lambda_k)^T, \sum_{k=1}^{k} A_k = 1, A_k \geq 0\}. \tag{3} \]

To characterize the efficiency of the units with respect to \( PPS_C \), we easily extend Definition 1 to deal with the newly defined production possibility set.

**Definition 2** (Cooper, Huang, Li, 2004) Let \((x_1, y_1), (x_2, y_2) \in PPS_C\) be two input-output pairs from the production possibility set \( PPS_C \). We say that \((x_1, y_1)\) dominates \((x_2, y_2)\) with respect to \( PPS_C \) if \(x_1 \leq x_2\) and \(y_1 \geq y_2\) with strict inequality holding for at least one of the components in the input or output vector. We say that \( DMU_0 \) is efficient with respect to \( PPS_C \) if there is no pair \((x, y) \in PPS_C\) dominating \( DMU_0 \) with respect to \( PPS_C \).

The output oriented BCC optimization model (Banker, Cooper, Charnes, 1978) is formulated as follows:

\[
\begin{align*}
\text{maximize } & \phi + \epsilon(\sum s_i^- + \sum s_j^+) \\
\text{subject to } & \sum_k x_{ik} \lambda_k + s_i^- = x_{i0} \text{ for all inputs } i = 1, ..., m, \\
& \sum_k y_{jk} \lambda_k - s_j^+ = \phi y_{j0} \text{ for all outputs } j = 1, ..., n, \\
& \sum_k \lambda_k = 1, \lambda_k \geq 0, s_i^-, s_j^+ \geq 0, \phi \text{ unconstrained,} \\
\end{align*}
\tag{4}
\]

where \( \epsilon \) is a non-Archimedean infinitesimal (a positive number smaller than any other positive number). The decision variables \( s_i^-, s_j^+ \) are again the slack and surplus variables for the inequalities \( X\lambda \leq x_0, Y\lambda \geq \phi y_0 \). The additional decision variable \( \phi \) represent the inefficiency factor of outputs of \( DMU_0 \). If \( \phi = 1 \), then \( DMU_0 \) with \((x_0, y_0)\) is sometimes called weakly efficient unit. Indeed, there could be more optimal solutions with \( \phi = 1 \) if we set \( \epsilon = 0 \); including the sum of slacks/surpluses into the objective, we select the “most dominant” among these solutions (c. f. the additive model), that is, with all slacks equal to zero. Non-Archimedean infinitesimal ensures that this sum of slacks and surpluses does not hurt the value of the objective function. We arrive at sufficient optimality conditions:

**Proposition 1** (Cooper, Huang, Li, 2004) If (a) \( \phi^* = 1 \), and (b) \( s^{-*} = s^{++} = 0 \), where \( * \) represents the optimum of (4), then \( DMU_0 \) is (fully) efficient with respect to \( PPS_C \).

Figure 2 demonstrates this situation: \( PPS_C \) is represented by the shaded region, its frontier is piecewise-linear line connecting units \( H, B, E, \) and \( H \). Points lying on the right-most part of the frontier (to the right of \( H \)) are solutions of (4) having \( \phi = 1 \), but the maximal slacks are not zero for these point—there is a point (namely \( H \)) which has the same value of the output but with smaller inputs, that is, dominating our considered points in sense of Definition 2.

**Figure 2** Example of the convex output oriented (BCC) DEA model

Source: Own processing of the data from Table 1.1 in Cooper, Seiford, Tone (2007)
For any inefficient unit, the non-zero elements of the variable $\lambda$ identify a point on the efficient frontier found as a convex combination of so-called peer units. Point $F$ on Figure 2 is not efficient (with $\phi < 1$): it is its only peer unit $E$ which has the optimal outputs with the same level of inputs.

The dual form of (4) is formulated by

$$\text{minimize } \sum_i u_i x_{io} + q \text{ subject to }$$
$$\Sigma_i u_i x_{ik} + q \geq \Sigma_j v_j y_{jk} \text{ for all units } k = 1, ..., K,$$
$$\Sigma_j v_j y_{jk} = 1 \text{ (dual constraint for } \phi),$$
$$u_i, v_j \geq \epsilon, q \text{ unconstrained},$$

which is the traditional formulation of BCC model (minimizing the input/output ratio). Non-Archimedean infinitesimal $\epsilon$ ensures that all inputs and outputs are considered in the analysis. The variable $q$, dual for the convexity constraint $\Sigma_k \lambda_k = 1$, is known as the variable return to scale factor.

### 3.3 Continuous (Linear) Production Possibility Set – CCR Model

Further relaxation of (4), removing the convexity constraint $\Sigma_k \lambda_k = 1$, results in the model given originally by Charnes, Cooper, Rhodes (1978) and known as CCR model. The production possibility set in this case is the convex cone containing observed units:

$$PPS_c = \{(x, y) | x = X\lambda, y = Y\lambda, \lambda = (\lambda_1, \lambda_2, ..., \lambda_k)^T, \lambda \geq 0 \}.$$  

To characterize the efficiency of the units with respect to $PPS_c$, we can use Definition 2, just replacing $PPS_c$ with $PPS_L$. The output oriented CCR optimization model (Charnes, Cooper, Rhodes, 1984) is formulated as follows:

$$\text{maximize } \phi + \epsilon(\Sigma_i s_i^- + \Sigma_j s_j^+ ) \text{ subject to }$$
$$\Sigma_k x_{ik}\lambda_k + s_i^- = x_{io} \text{ for all inputs } i = 1, ..., m,$$
$$\Sigma_k y_{jk}\lambda_k - s_j^+ = \phi y_{jo} \text{ for all outputs } j = 1, ..., n,$$
$$\lambda_k \geq 0, s_i^-, s_j^+ \geq 0, \phi \text{ unconstrained},$$

The interpretations of the model and optimality conditions are analogous to the convex case. The situation is demonstrated on Figure 3; you can notice the production possibility set in conic form.

**Figure 3** Example of the linear output-oriented CCR DEA model

Source: Own processing of the data from Table 1.1 in Cooper, Seiford, Tone (2007)

The dual form of (7) is formulated by

$$\text{minimize } \sum_i u_i x_{io} \text{ subject to }$$
$$\Sigma_i u_i x_{ik} \geq \Sigma_j v_j y_{jk} \text{ for all units } k = 1, ..., K,$$
$$\Sigma_j v_j y_{jk} = 1 \text{ (dual constraint for } \phi),$$
$$u_i, v_j \geq \epsilon.$$
The missing variable $q$ (i.e., $q = 0$) is manifestation of the property of constant return to scale.

4 Stochastic Extensions to Production Possibility Sets

The data represented by the matrices $X$, $Y$ are usually not deterministic. We will now consider the inputs and outputs to be random variables. Hence, the matrices $X$, $Y$ and corresponding PPS are also random. To deal with resulting random constraints, we also define a (sufficiently small) tolerance or risk level $\alpha \in (0; 1)$. It is worth to note here that the notion of stochastic efficiency is not uniquely determined but adapted to a particular situation.

4.1 Stochastic additive model

In the case of discrete production possibility set defined by (1).

Definition 3 (Cooper et al., 1988) We say that $DMU_0$ is not stochastically dominates with respect to $PPS_j$ if for each $\lambda \in \{0, 1\}$ with $\sum \lambda = 1$ we have

$$\mathbb{P}\{X\lambda \leq x_0, Y\lambda \geq y_0\} \leq \alpha.$$  

(9)

The stochastic extension of the additive 0–1 model reads simply as

$$\beta^* = \max \mathbb{P}\{X\lambda \leq x_0, Y\lambda \geq y_0\}. \quad (10)$$

It is now easy to see the optimality condition for stochastic efficiency in this case:

Proposition 2 (Cooper, Huang, Li, 2004) $DMU_0$ is $\alpha$-stochastically efficient with respect to $PPS_j$ if and only if $\beta^* > \alpha$.

The concept of stochastic efficiency is illustrated on Figure 4. Conclusions about efficiency must be made with respect to uncertain position of the input-output points: for example, units $B$ and $C$ can exchange their positions so that $C$ is dominating $B$ for a particular instance. In stochastic terms, both $B$ and $C$ must be indicated as efficient if the probability of such event is sufficiently high. The only remaining non-efficient points are the units $F$ and $G$ for which this probability is small.

Figure 4 Example of the stochastic additive DEA model

Source: Own processing of the data from Table 1.1 in Cooper, Seiford, Tone (2007)

4.2 Stochastic continuous model

Continuous extension to additive stochastic dominance is straightforward:

Definition 4 (Cooper et al., 1988) We say that $DMU_0$ is $\alpha$-stochastically efficient with respect to $PPS_j$ if for each $\lambda \geq 0$ with $\sum \lambda = 1$ we have

$$\mathbb{P}\{X\lambda \leq x_0, Y\lambda \geq y_0\} \leq \alpha.$$  

(11)

Applying necessary and sufficient conditions for stochastic dominance (see Cooper et al., 1988), the stochastic extension to BCC takes the following form known as almost 100% confidence chance-constrained problem
\[
\beta^* := \max \{ \mathbb{P}\left[ \sum_i s_i^- + \sum_f s_f^+ < 0 \right] \} \text{ subject to }
\begin{align*}
\mathbb{P}\left[ \sum_k x_{ik} \lambda_k + s_i^- < x_{i0} \right] & \geq 1 - \epsilon \quad \text{for all inputs } i = 1, \ldots, m, \\
\mathbb{P}\left[ \sum_k y_{jk} \lambda_k - s_j^+ > y_{j0} \right] & \geq 1 - \epsilon \quad \text{for all outputs } j = 1, \ldots, n, \\
\sum_k \lambda_k &= 1, \lambda_k \geq 0, s_i^-, s_j^+ \geq 0.
\end{align*}
\] (12)

Proposition 3 (Cooper et al., 1988) If \(DMU_0\) is \(\alpha\)-stochastically efficient, then \(\beta^* \leq \alpha\). If \(\beta^* > \alpha\) then \(DMU_0\) is not \(\alpha\)-stochastically efficient.

4.3 Marginal chance-constrained model

Relaxing the almost 100% (namely \(1 - \epsilon\)) confidence, Cooper et al. (2002, 2003) introduced a marginal chance-constrained optimization model for BCC of the form

\[
\begin{align*}
\text{maximize } & \phi \\
\text{subject to } & \mathbb{P}\left[ \sum_k x_{ik} \lambda_k + s_i^- = x_{i0} \right] \geq 1 - \alpha \quad \text{for all inputs } i = 1, \ldots, m, \\
& \mathbb{P}\left[ \sum_k y_{jk} \lambda_k - s_j^+ = y_{j0} \right] \geq 1 - \alpha \quad \text{for all outputs } j = 1, \ldots, n, \\
& \sum_k \lambda_k = 1, \lambda_k \geq 0, s_i^-, s_j^+ \geq 0.
\end{align*}
\] (13)

Problem (13) implies the following definition of marginal stochastic efficiency:

Definition 5 (Cooper et al., 2002) We say that \(DMU_0\) is marginally \(\alpha\)-stochastically efficient if (a) \(\phi^* = 1\), and (b) \(s^* = s^{**}\) for all (alternate) optimal solutions, where * represents the optimum of (13).

Problem (13) is a special case of the standard individual chance-constrained problem. If the constraint rows follow normal distribution, (13) reduces to a quadratic optimization problem (involving normal quantiles \(\Phi^{-1}(\alpha)\)). We refer to Prékopa (2003) to a survey of solution methods for such problems; see also Cheng, Houda, Lisser (2015) for a recent contribution (involving dependency properties and second-order cone programming methods).

5 Conclusions

In the paper, we presented several types of deterministic and stochastic models. The non-classical way of the construction of deterministic models – based on domination of units – allows us to show a close connection between deterministic and stochastic models. We also presented that these deterministic models coincide with classical ones, because they are in fact their dual forms.

References


