SOME PROPERTIES OF GEOMETRIC DEA MODELS

Ozren Despić
Aston Business School, Aston University,
Birmingham, UK
E-mail: o.despic@aston.ac.uk

Abstract

Some specific geometric data envelopment analysis (DEA) models are well known to the researchers in DEA through so-called multiplicative or log-linear efficiency models. Valuable properties of these models were noted by several authors but the models still remain somewhat obscure and rarely used in practice. The purpose of this paper is to show from a mathematical perspective where the geometric DEA fits in relation to the classical DEA, and to provide a brief overview of some benefits in using geometric DEA in practice of decision making and/or efficiency measurement.

Key words: *Data envelopment analysis, Technical efficiency, Weighted geometric mean*

1. INTRODUCTION

One of the main concepts addressed by DEA is the concept of technical efficiency, which, in simplest terms, can be defined as a relative measure of the success of a Decision Making Unit (DMU) in maximizing its desirable outputs while at the same time minimizing its relevant inputs. To make this definition practical in terms of measuring and analyzing efficiency, it is necessary to construct some kind of a function whose arguments will be all the relevant inputs and all the desirable outputs and which has to satisfy two basic properties: to be directly proportional to all the outputs and inversely proportional to all the inputs. Clearly, one could come up with infinitely many different functions satisfying these two basic properties. To narrow down the set of possible functions, we need to impose further properties that such a function should satisfy until we finally obtain a workable formulation that can be used to evaluate the efficiency of a DMU. But before we do, let us see a simple general functional form for efficiency measure e_k defined for DMU k:

$$e_k = \frac{f_k}{f_{\max}} = \frac{f(Input_{1k}, Input_{2k}, \ldots, Input_{mk}, Output_{1k}, Output_{2k}, \ldots, Output_{sk})}{f_{\max}} \tag{1}$$

In case of a single relevant input and a single desirable output, the function f is customarily defined in the form of a ratio:

$$f = \frac{Output_k}{Input_k} \tag{2}$$

This function satisfies the basic two properties mentioned earlier and its value tells us how much of output has been attained per one unit of input. Organizations, such as business firms, hospitals, educational institutions, etc., are frequently using the above ratio form to evaluate productivity of its units. Some examples are “sales per salesperson hour”, “inpatients per doctor employed” or “number of publications per faculty member”. The measures like these are also known as partial productivity measures. The word partial is used here since these measures do not capture productivity based on all desirable outputs and all relevant inputs but deal only with one input-output pair. Total factor productivity measure is what we would like to name the value of function f in general case and this could be some kind of output-to-input ratio value where all desirable outputs and all relevant inputs are included. To achieve this, we need to generalize the above ratio for the case when there is more than one input and more than one output. One of the most frequently presented generalizations is:

$$f = \frac{\sum_j b_j y_j}{\sum_i a_i x_i}, \tag{3}$$

where a_i and b_j are the weights applied to input x_i and output y_j, respectively. These weights could not be negative since otherwise the basic two properties of function f, mentioned earlier, would not be respected.

The basic idea behind classical DEA is to derive efficiency measure e_k of DMU_k under the following two conditions:

1. The parameters of function f are not to be specified in advance; instead, they are left to be determined by each unit k being assessed so that the selected parameters maximise its efficiency score. This is why DEA belongs to the group of nonparametric approaches to efficiency measurement and this is also why DEA is said to satisfy strict equity criteria.
2. The value of f_{max} in (1) must be obtained using the same function f as for unit under assessment but applied to the inputs and outputs of an observed unit. The observed unit normally selected for this purpose is the one which will maximise the value of function f using the parameters selected by unit k. Note that the observed unit taken for this purpose can also be the unit k itself. For this reason, DEA efficiency scores are bound from above by 1. Due to this, it is said that DEA focuses on revealed best-practice frontier.

When the form of function f is specified as in (3), then we have a seed for many classical DEA model. Function f in (3) is truly a generalization of (2) since for the single input - single output case they both yield the same values for e_k in (1). However, the generalization of (2), as shown in (3) is not the only possible generalization. Other possible generalisations will be explored in the next section. One of them will be the seed for all the geometric DEA models. For now, let us just observe the formulation for the efficiency of DMU$_k$ when the above conditions are applied to the single input single output case:

$$e_k = \frac{f_k}{f_{\max}} = \frac{f(X_k, Y_k)}{f_{\max}} = \frac{Y_k / X_k}{\max\limits_p \left(Y_p / X_p \right)} = \min_p \frac{Y_k / X_k}{Y_p / X_p} = \min_p \frac{X_p / X_k}{Y_p / Y_k} = \min_p \left(\frac{X_p}{X_k} \times \frac{Y_k}{Y_p} \right),$$

where unit p is selected among all the observed units.

It is instructive to note that the last formulation for e_k in the above line of equalities can be interpreted in the following way: efficiency of DMU$_k$ is equal to the product between its input factor efficiency (X_p/X_k) and its output factor efficiency (Y_k/Y_p), where the factor efficiencies are obtained with respect to a unit p which will minimise the product of two factor efficiencies.

2. MATHEMATICAL INTRODUCTION TO EFFICIENCY

In this section, we will examine alternative forms of function f that could be used in (1) while satisfying the property that, for the single input – single output case, it yields the same efficiency as (2) when used in (1).

An average value of a set of positive real numbers, $a_1, a_2, \ldots, a_n$, may be defined in a number of ways. Some of the common definitions include:

arithmetic mean: $A = A(a_1, a_2, \cdots, a_n) = \dfrac{a_1 + a_2 + \cdots + a_n}{n}$,

geometric mean: $G = G(a_1, a_2, \cdots, a_n) = \sqrt[n]{a_1 a_2 \cdots a_n}$,

harmonic mean: $H = H(a_1, a_2, \cdots, a_n) = \dfrac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}}$.

It is well known that $H \le G \le A$ and both inequalities become equalities if and only if $a_1 = a_2 = \ldots = a_n$.

These definitions are usually adequate in applications where the underlying data are of equal importance. In some cases, however, the relative importance of the data is essential, and may be expressed numerically, in the form of non-negative real values: $w_1, w_2, \ldots, w_n$, called weights. The weights can be normalized so that $w_1 + w_2 + \ldots + w_n = 1$. Under such circumstances, for a given weight vector $w = (w_1, w_2, \ldots, w_n)$, the *weighted arithmetic mean* is defined as the value

$$A_w = A_w(a_1, a_2, \cdots, a_n) = w_1 a_1 + w_2 a_2 + \cdots + w_n a_n .$$

Similarly, the *weighted geometric mean* is:

$$G_w = G_w\left(a_1, a_2, \cdots, a_n\right) = a_1^{w_1} a_2^{w_2} \cdots a_n^{w_n},$$

and the *weighted harmonic mean* is:

$$H_w = H_w\left(a_1, a_2, \cdots, a_n\right) = \frac{1}{\dfrac{w_1}{a_1} + \dfrac{w_2}{a_2} + \cdots + \dfrac{w_n}{a_n}}.$$

Note that if the underlying data is of equal importance, then $w = (1/n, 1/n, \ldots, 1/n)$, and so $A_w = A$, $G_w = G$ and $H_w = H$. It is true, in general, that

$$H_w \le G_w \le A_w \tag{4}$$

Both inequalities become equalities if and only if $a_1 = a_2 = \ldots = a_n$.

Let $T = (t_{ip})_{m\times n}$ be a matrix with positive entries, and let W_m be the set of all vectors $w = (w_1, w_2, \ldots, w_n)$, in $\Re^m$ such that $w_1 + w_2 + \ldots + w_n = 1$ and $w_i \ge 0$ ($i = 1, 2, \ldots, m$). A *convex linear combination* of the row vectors of T is a vector of the form wT, where $w \in W_m$. Note that the coordinates of wT are simply the weighted arithmetic means of the corresponding coordinates of the row vectors of T. Therefore we shall adopt the following notation:

$$A_w(T) = wT = \left(A_w\left(t_{1p}, t_{2p}, \cdots, t_{mp}\right)\right)_{1\times n}.$$

If the weighted arithmetic means are replaced by the corresponding weighted geometric means, we shall denote the resulting vector as follows:

$$G_w(T) = \left(G_w\left(t_{1p}, t_{2p}, \cdots, t_{mp}\right)\right)_{1\times n}.$$

Similarly, by taking the weighted harmonic means, coordinatewise over the row vectors of the matrix T, we shall denote the resulting vector as follows:

$$H_w(T) = \left(H_w\left(t_{1p}, t_{2p}, \cdots, t_{mp}\right)\right)_{1\times n}.$$

Note that, in view of the inequalities (4), the following vector inequalities must hold:

$$H_w\left(T\right) \le G_w\left(T\right) \le A_w\left(T\right). \tag{5}$$

Let us now consider n DMUs, each utilizing m inputs and generating s outputs. Let the input values be represented by the matrix $X = (x_{ip})_{m\times n}$, and let the output values be represented by the matrix $Y = (y_{jp})_{m\times n}$. As a matter of convenience, we shall assume that all entries in the matrices X and Y are positive. In practice, this may be accomplished by replacing all zero entries with a sufficiently small positive value ε.

As per conditions specified in the previous section, standard measure for the efficiency of unit k ($1 \le k \le n$) may be expressed as:

$$e_k = \max_{a_i, b_j \ge 0} \min_p \frac{\sum_{j=1}^{s} b_j y_{jk}}{\sum_{i=1}^{m} a_i x_{ik}} \Bigg/ \frac{\sum_{j=1}^{s} b_j y_{jp}}{\sum_{i=1}^{m} a_i x_{ip}}. \tag{6}$$

The above measure of efficiency was first defined by Charnes et al (1978). It is normally denoted as *CCR* efficiency, with the acronym CCR referring to authors' names. In order to arrive at a more convenient expression for the efficiency measure of the k-th DMU, let us transform the matrices X and Y as follows. Let $X(k)$ be the matrix obtained from X by replacing each entry x_{ip} with x_{ip}/x_{ik}. In particular, the k-th column of the matrix $X(k)$ is a column of ones. Similarly, let $Y(k)$ be the matrix obtained from Y by replacing each entry y_{jp} with y_{jk}/y_{jp}. Thus the k-th column of the matrix $Y(k)$ is also a column of ones. Furthermore, for any vectors $u = (u_1, u_2, \dots, u_n)$, and $v = (v_1, v_2, \dots, v_n)$, $v_p \ne 0$ ($p = 1, 2, \dots, n$), let us write

$$\left[u : v\right] = \left(\frac{u_1}{v_1}, \frac{u_2}{v_2}, \cdots, \frac{u_n}{v_n}\right).$$

Therefore, if we denote the column vectors of the matrix X by X_p, then the corresponding column vectors of the matrix $X(k)$ are $[X_p : X_k]$, ($p = 1, 2, \dots, n$). Similarly, if we denote the column vectors of the matrix Y by Y_p, then the corresponding column vectors of the matrix $Y(k)$ are $[Y_k : Y_p]$, ($p = 1, 2, \dots, n$). Note that the elements of matrix $X(k)$ can be interpreted as relative input strengths or input factor efficiencies of unit k with respect to all other units. Similarly, the elements of matrix $Y(k)$ can be interpreted as relative output strengths or output factor efficiencies of unit k with respect to all other units. For example, take the value located at i-th row and p-th column of matrix $X(k)$. If this value is greater than 1 then unit k is doing better than unit p with respect to input i (i.e, unit k is using less of input i than unit p). Similarly, for a value located at j-th row and p-th column of matrix $Y(k)$: if the value is greater than 1 then unit k is doing better than unit p with respect to output j. Values smaller than 1 would clearly indicate the opposite.

According to Theorem 1 in Despic et al. (2007), (6) is equivalent to:

$$CCR: \quad e_k = \max_{\substack{\sum_j b_j = 1, b_j \ge 0 \\ \sum_i a_j = 1, a_i \ge 0}} \min_p \frac{\sum_i a_i \dfrac{x_{ip}}{x_{ik}}}{\sum_j b_j \dfrac{y_{jp}}{y_{jk}}}. \tag{7}$$

Now, in view of the notation introduced above, (7) can be re-written as:

$$CCR: \quad e_k = \max_{\substack{b \in W_s \\ a \in W_m}} \min \left[A_a \left(X\left(k\right) \right) \otimes H_b \left(Y\left(k\right) \right) \right], \tag{8}$$

Symbol ⊗ is used to denote coordinate-wise product of two vectors.

Another measure for the efficiency, called the *harmonic efficiency* (*HE*), was also introduced in Despic et al. (2007) as

$$HE: \quad \overline{e}_k = \max_{\substack{\sum_j b_j = 1, b_j \geq 0 \\ \sum_i a_j = 1, a_i \geq 0}} \min_p \frac{1}{\sum_j b_j \frac{y_{jp}}{y_{jk}} \times \sum_i a_i \frac{x_{ik}}{x_{ip}}}. \tag{9}$$

In view of the notation introduced above, (9) can be re-written as:

$$HE: \quad \overline{e}_k = \max_{\substack{b \in W_s \\ a \in W_m}} \min \left[H_a \left(X\left(k\right) \right) \otimes H_b \left(Y\left(k\right) \right) \right]. \tag{10}$$

We shall now compare the expressions on the right sides of (8) and (10). By the inequalities in (5), the following vector inequality must hold:

$$H_a \left(X\left(k\right) \right) \leq A_a \left(X\left(k\right) \right).$$

Hence $\left[H_a \left(X\left(k\right) \right) \otimes H_b \left(Y\left(k\right) \right) \right] \leq \left[A_a \left(X\left(k\right) \right) \otimes H_b \left(Y\left(k\right) \right) \right]$,

and so, by (8) and (10),

$$\overline{e}_k \leq e_k \quad \text{or} \quad HE \leq CCR. \tag{11}$$

A third type of a measure for the efficiency of the k-th DMU, called the *DEA-R efficiency*, was introduced in Despic et al. (2007) as the standard efficiency applied to a derived set of input-output data. Specifically, the new inputs are represented by the $1 \times n$ matrix I whose entries are all ones, while the new outputs are represented by the (sm)$\times n$ matrix R, whose entries are all the possible ratios $r_{(i,j)p} = y_{jp}/x_{ip}$, where $1 \leq i \leq m$, $1 \leq j \leq s$, and $1 \leq p \leq n$. Each pair of indices (i,j) determines a row of the matrix R as the vector $[Y_j : X_i]$, where Y_j and X_i are the j-th row of the matrix Y and the i-th row of the matrix X, respectively. Let us note that the order in which the *sm* rows of the matrix R are arranged is irrelevant, since the resulting formula for the *DEA-R* efficiency will be the same:

$$\hat{e}_k = \max_{\substack{\sum_{(i,j)} c_{(i,j)} = 1 \\ c_{(i,j)} \geq 0}} \min_p \frac{1}{\sum_{(i,j)} c_{(i,j)} \frac{y_{jp}}{y_{jk}} \frac{x_{ik}}{x_{ip}}}. \tag{12}$$

In view of the notation introduced above, (12) can be called *harmonic ratio efficiency* and re-written as

$$HRE: \quad \hat{e}_k = \max_{c \in W_{sm}} \min \left[H_c \left(R\left(k\right) \right) \right]. \tag{13}$$

We shall now compare the expressions on the right sides of (10) and (13). Given any weight vector $b = (b_1, b_2, \ldots, b_n)$ in W_s and any weight vector $a = (a_1, a_2, \ldots, a_n)$ in W_m, let us define a vector c in $\Re^{sm}$ by setting $c_{(i,j)} = a_i b_j$ for any pair (i,j) with $1 \le i \le m$ and $1 \le j \le s$. Then

$$\sum\nolimits_{(i,j)} c_{(i,j)} = \sum\nolimits_{(i,j)} a_i b_j = \left(\sum\nolimits_i a_i\right)\left(\sum\nolimits_j b_j\right) = 1.$$

Therefore c is a weight vector in W_{sm}. Furthermore,

$$\frac{c_{(i,j)}}{r_{(i,j)k} / r_{(i,j)p}} = \frac{a_i b_j}{\left(y_{jk}/x_{ik}\right)/\left(y_{jp}/x_{ip}\right)} = \left(\frac{a_i}{x_{ip}/x_{ik}}\right)\left(\frac{b_j}{y_{jk}/y_{jp}}\right).$$

By taking the summation over all indices (i,j), with $1 \le i \le m$ and $1 \le j \le s$, we obtain

$$H_c\left(R(k)\right) = \left[H_a\left(X(k)\right) \otimes H_b\left(Y(k)\right)\right].$$

It follows that the maximum taken in (10) is over a subset of the values whose maximum is taken in (13). Hence

$$\bar{e}_k \le \hat{e}_k \quad \text{or} \quad HE \le HRE. \tag{14}$$

Let us observe that each of our measures for the efficiency of the k-th DMU has been expressed via suitable weighted arithmetic and/or weighted harmonic means. Also, they all produce the same efficiency score in single input – single output case, since they are all proper generalisation of (2) when used in (1) for multiple input – multiple output case.

We now want to introduce a new type of a measure for the efficiency of the k-th DMU, based on the weighted geometric means. We shall define the efficiency measure $\tilde{e}_k$ by replacing the weighted harmonic mean in the equation (13) with the corresponding weighted geometric mean. Thus, *geometric ratio efficiency* (*GRE*) is defined as

$$\tilde{e}_k = \max_{c \in W_{sm}} \min\left[G_c\left(R(k)\right)\right]. \tag{15}$$

By (5), it follows that

$$\hat{e}_k \le \tilde{e}_k \quad \text{or} \quad HRE \le GRE. \tag{16}$$

It is interesting to note that there is another way to arrive at the same definition for $\tilde{e}_k$. Namely, if we replace both weighted means in the equation (8) with the corresponding weighted geometric means, we may define *geometric efficiency* measure $\breve{e}_k$ by:

$$GE: \quad \breve{e}_k = \max_{\substack{b \in W_s \\ a \in W_m}} \min\left[G_a\left(X(k)\right) \otimes G_b\left(Y(k)\right)\right]. \tag{17}$$

By (5), we know that $H_a\left(X(k)\right) \le G_a\left(X(k)\right)$, and $H_b\left(Y(k)\right) \le G_b\left(Y(k)\right)$.

Hence $\left[H_a\left(X(k)\right) \otimes H_b\left(Y(k)\right)\right] \le \left[G_a\left(X(k)\right) \otimes G_b\left(Y(k)\right)\right]$.

By (10) and (17), it follows that

$$\overline{e}_k \le \breve{e}_k \quad \text{or} \quad HE \le GE\,. \tag{18}$$

It is relatively straightforward to show that geometric efficiency *GE* is equivalent to geometric ratio efficiency *GRE*. This equivalency is particularly important in applications where we may want to impose some restriction on some specific pairs of inputs and outputs in terms of their importance and relative contribution to efficiency. If some specific input-output pairs are considered as not meaningful, then we can switch from *GE* model to *GRE* model and exclude the ratios corresponding to those input-output pairs.

In an analogous way to (10) and (13), it is possible to define *arithmetic efficiency* (*AE*), and *arithmetic ratio efficiency* (*ARE*). When all these different efficiencies are compared using (5), we have the following relationships:

$$HE \le HRE \le GE = GRE \le AE \le ARE\,. \tag{19}$$

As for the standard *CCR* efficiency, we know that it is never smaller than *HE* and never larger than *AE*. Hence, in addition to (19), we have:

$$HE \quad \le \quad CCR \quad \le \quad AE\,. \tag{20}$$

With (19) and (20), we have effectively specified the ordering relationships (in terms of the efficiency scores produced) among all those models, each of which represents a different generalisation of (2) when used in (1) for multiple input – multiple output case.

3. DEA EFFICIENCY MEASURES AS AGGREGATION OPERATORS

Efficiency measures of unit *k* obtained using *CCR*, *HE*, *GE* or *AE* models can be seen as values obtained using different aggregation operators applied on the same data set, which is made of two matrices: one representing a collection of relative input strengths of unit *k*, *X*(*k*), and the other one representing collection of relative output strengths of unit *k, Y*(*k*). We can observe that all four models are using some specific case of a weighted power mean to aggregate columns of both matrices. Focusing on one and the same column in both matrices, it can be stated that each of the four models yields a single measure for unit *k* by forming the product between a weighted power mean taken over its relative input strengths and a weighted power mean taken over its relative output strengths. The questions we want to consider in this section are: “What is the role of the product between the two means?” and “Is there any obvious advantage or disadvantage in using a specific weighted mean to aggregate relative strengths of unit *k*?”

Before we consider the above questions, let us just take a brief look at the full matrix of relative strengths *RS*(*k*) formed by putting together matrices *Y*(*k*) and *X*(*k*).

$$RS(k)=\begin{bmatrix} y_{1k}/y_{11} & y_{1k}/y_{12} & \cdots & y_{1k}/y_{1p} & \cdots & 1 & \cdots & y_{1k}/y_{1n} \\ y_{2k}/y_{21} & y_{2k}/y_{22} & \cdots & y_{2k}/y_{2p} & \cdots & 1 & \cdots & y_{2k}/y_{2n} \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ y_{sk}/y_{s1} & y_{sk}/y_{s2} & \cdots & y_{sk}/y_{sp} & \cdots & 1 & \cdots & y_{sk}/y_{sn} \\ x_{11}/x_{1k} & x_{12}/x_{1k} & \cdots & x_{1p}/x_{1k} & \cdots & 1 & \cdots & x_{1n}/x_{1k} \\ x_{21}/x_{2k} & x_{22}/x_{2k} & \cdots & x_{2p}/x_{2k} & \cdots & 1 & \cdots & x_{2n}/x_{2k} \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ x_{m1}/y_{mk} & x_{m2}/y_{mk} & \cdots & x_{mp}/y_{mk} & \cdots & 1 & \cdots & x_{mn}/y_{mk} \end{bmatrix}$$

Each column of matrix *RS*(*k*) contains *m*+*s* relative strengths of unit *k*; column 1 shows its component strengths with respect to unit 1, column 2 with respect to unit 2 and so on. All the four models, *CCR*, *HE*, *GE* or *AE* use the same DEA-like approach of choosing the relevant column(s) and the best set of weights to aggregate *m*+*s* factors into a single one, which is related to the max-min part of the models. This is based on a familiar idea of a game played between the unit assessed and the assessor. The unit is allowed to choose the weights for aggregating its component strengths while the assessor picks the column with the smallest score to be the final efficiency score of the unit. The only thing different across the four models is the functional form of the aggregation operator used, which is something rarely discussed and questioned in classical DEA analysis.

Representing all the four models through weighted power mean, as in (21),

$$PE: \quad E_k = \max_{\substack{b \in W_s \\ a \in W_m}} \min \left[P_a\left(X(k)\right) \otimes P_b\left(Y(k)\right) \right], \tag{21}$$

we can see that the product between the two weighted power means is a common feature for all four models. Since the final score is the product between two values, this means that the two values represent aggregated scores of two strongly separated categories (in production context these categories are input and outputs), each of which is equally valued in terms of its contribution to the final score (since $\sum_i a_i = \sum_j b_j = 1$). Strong separation between the two categories means that higher values of the relative strengths found in one category cannot compensate for lower values of the relative strengths found in the other category. Observing this property from a general multi-criteria decision making framework, it should not be difficult to see how restrictive this property could be. While within certain context it may be quite natural to strongly separate a set of criteria into two categories (so to prevent any substitutability of factors across the categories and allow for it within the categories), it is also possible to have the case where this separation is either not needed at all or where we need to split the set of criteria into more than two strongly separated categories. In addition, the requirement to make the categories equally valued (equally important) is certainly overly restrictive in general case. Hence, splitting criteria into two categories and forming the product between the two will provide a proper model only for a special case when there is natural separation of the criteria into two categories (based on considerations of substitutability among the criteria) and when the two categories are equally important with respect to the final measure. Clearly, any model that can avoid these restrictions should be preferred in practice.

The way individual relative strengths are aggregated within a single category is different for different DEA models. *CCR* model takes weighted harmonic mean within one category and weighted arithmetic mean within the other category. *HE*, *GE*, and *AE* use weighted harmonic, weighted geometric and weighted arithmetic mean within both categories, respectively. From a mathematical point of view, it is difficult to justify the use of different means to aggregate relative strengths within different categories. Within production context this perhaps makes sense due to the fact that the final measure is perceived as the ratio between total virtual output and total virtual input. However, in general context, and bearing in mind that the relative strengths formed are dimensionless index-like values, then there is no obvious reason why would any weighted mean be preferred over the other. Still, if some special cases are considered then the weighted geometric mean is the only one which does not violate some desirable properties of the model. To illustrate this, consider a special case where there happens to be a mutual agreement about the set of weights to be applied to all n units (i.e., not allowing any variability in weights for different units). Using any weighted mean to aggregate relative strengths column-wise, we can obtain n different scores for each of n different *RS* matrices (there is one *RS* matrix for each unit). Since the same set of weights is used within all n matrices, it would be natural to expect that the rank ordering of n scores obtained from one *RS* matrix remain the same for all the other *RS* matrices. Unfortunately, this property is not preserved by any other weighted mean but the weighted geometric mean. The property of the weighted geometric mean to preserve this ordering is closely related to the similar property of being the resistant to rank reversal problem in Analytic Hierarchy Process (AHP). In fact, as we will see in the next section, the use of weighted geometric mean in (21) makes it possible to create a much more flexible DEA-like model, which is at the same time nothing else but a generalisation of the multiplicative AHP model.

4. GEOMETRIC DEA MODELS AND THEIR PROPERTIES

The *GE* model, as defined in (17), can be seen as a seed for all other models in geometric DEA, much like the *CCR* model can be seen as a seed for all other classical DEA models. To start with, let us consider linear formulation of the *GE* and *CCR* model as well as some of its variations and compare the geometric DEA models with their counterparts in the set of classical DEA models.

The *GE* model in (17) can be transformed into linear programming problem using the following transformations:

$$\breve{e}_k = \max_{\substack{b\in W_s \\ a\in W_m}} \left\{ \omega_k \;\middle|\; \prod_i \left(\frac{x_{ip}}{x_{ik}}\right)^{a_i} \prod_j \left(\frac{y_{jk}}{y_{jp}}\right)^{b_j} \ge \omega_k \right\} = \max_{\substack{b\in W_s \\ a\in W_m}} \left\{ \omega_k \;\middle|\; \omega_k \frac{\prod_j \left(\frac{y_{jp}}{y_{jk}}\right)^{b_j}}{\prod_i \left(\frac{x_{ip}}{x_{ik}}\right)^{a_i}} \le 1 \right\}$$

After taking the log of the last formulation, the following linear programming formulation for *GE* model is obtained:

$$Max\ \theta_k$$

s.t.

$$\sum_j b_j \ln\left(\frac{y_{jp}}{y_{jk}}\right) - \sum_i a_i \ln\left(\frac{x_{ip}}{x_{ik}}\right) + \theta_k \le 0 \quad \forall p, \tag{22}$$

$$\sum_i a_i = 1, \quad a_i \ge 0$$

$$\sum_j b_j = 1, \quad b_j \ge 0$$

where $\theta_k = \ln(\omega_k)$. Efficiency score ω_k is obtained by solving the model in (22), which yields optimal value for θ, which is then used to calculate efficiency score $\omega_k = exp(\theta_k)$.

Following similar transformations, *CCR* model in (7) can be transformed as follows:

$$e_k = \max_{\substack{b\in W_s \\ a\in W_m}} \left\{ \omega_k \;\middle|\; \frac{\sum_i a_i \frac{x_{ip}}{x_{ik}}}{\sum_j b_j \frac{y_{jp}}{y_{jk}}} \ge \omega_k \right\} = \max_{\substack{b\in W_s \\ a\in W_m}} \left\{ \omega_k \;\middle|\; \omega_k \left(\sum_j b_j \frac{y_{jp}}{y_{jk}}\right) - \left(\sum_i a_i \frac{x_{ip}}{x_{ik}}\right) \le 0 \right\}$$

The last expression can be then converted into the following linear programming model:

$$Max\ \omega_k$$

s.t.

$$\sum_j \beta_j \left(\frac{y_{jp}}{y_{jk}}\right) - \sum_i a_i \left(\frac{x_{ip}}{x_{ik}}\right) \le 0 \quad \forall p, \tag{23}$$

$$\sum_i a_i = 1, \quad a_i \ge 0$$

$$\sum_j \beta_j = \omega_k, \quad \beta_j \ge 0$$

where $\beta_j = \omega_k \times b_j$.

The linear programming formulation of the *CCR* model, as shown in (23), is not frequently seen in literature. However, it can be obtained directly from the classical *CCR* input-oriented envelopment formulation, first by dividing all input related constraints by x_{ik} and all output related constraints by y_{jk}. Taking the dual form of such a transformed problem would lead us directly to the form shown in (23).

There are several things worth noting when comparing the models (22) and (23). First, it is important to understand that *CCR*, given in (7), could be inverted and reformulated as

$$\frac{1}{e_k} = \min_{\substack{\sum_j b_j = 1, b_j \geq 0 \\ \sum_i a_j = 1, a_i \geq 0}} \max_p \frac{\sum_j b_j \frac{y_{jp}}{y_{jk}}}{\sum_i a_i \frac{x_{ip}}{x_{ik}}}.$$

Using the same transformation steps as in the process of obtaining (22) and (23), the above formulation would lead us to the *CCR* output-oriented model similar to the one in (23). The main difference would be that in the output-oriented model we would be minimising ω_k, subject to the same set of constraints but with $\Sigma\alpha_i = \omega_k$ and $\Sigma b_j = 1$, and where $\omega_k = 1/e_k$ and $\alpha_i = \omega_k \times a_j$. Inverting the *GE* formulation, on the other hand, yields the model which is identical to (22). In other words the optimal values for weights a_i and b_j would be the same for both models. This is because the *GE* model essentially treats inputs as inverted outputs and/or outputs as inverted inputs. If, for example, we invert all the inputs and treat them as outputs, but still keeping them in a separate group from the original set of outputs, we would then only need to change minus signs in (22) into plus signs and the results obtained would be identical to the results we had before. *CCR* model, on the other hand, does not offer any foreseeable way of converting inputs into outputs or outputs into inputs without making changes to the optimal solutions.

As we will see later, this property of the *GE* model to treat inputs as inverted outputs will be very convenient in formulating a geometric DEA model when faced with multiple categories and multiple levels. This will essentially enable us to easily deal with any hierarchical structure and not only the standard one level – two categories structure that fits the classical division of factors into a set of inputs and set of outputs.

Before we present some important variations of the models in (22) and (23), it will be very useful to better understand the weights in these models. The weights a_i and b_j in model (22) and a_i and β_j in model (23) have similar interpretation and they are directly related to what is known as virtual inputs and virtual outputs in classical DEA. Looking at model (23), we can see that the input weights add up to 1 while the output weights add up to the efficiency score of unit *k*, just as the virtual inputs add up to 1 and virtual outputs add up to the efficiency score of unit *k* in the standard multiplier formulation of the *CCR* input-oriented model. The weights in (22) and (23) are dimensionless and hence, just like virtual inputs and outputs in classical DEA, they reveal the relative contribution of each input and output to the efficiency score of the unit assessed. Clearly, the true interpretation is not quite as simple. It is somewhat simpler in case of the *GE* model (22) since both sets of weights add up to 1. The sum to unity is convenient to have since the value of each weight can really be treated as a true proportion of the contribution of the corresponding factor to the efficiency score of the unit assessed. But, what are the factors to which the weights are attached? They are not simply inputs and outputs of unit k. They are relative input and output values compared to the input and output values of another unit. Earlier, we called these ratios relative input strengths and relative output strengths of the unit under assessment. However, it is important to realise that the weights will be determined only when the relative input and relative output strengths are formed with respect to unit(s) from the best practice frontier. Those best practice units will be picked from that portion of the frontier which represents the set of production plans that are most similar to the current production plan of the unit assessed (similar in terms of relative intensities of inputs and outputs). In other words, relative strengths of inputs and outputs can be seen as relative values taken with respect to some ideal values and where the ideal values come from the observed best practices. Hence, the weights can be understood as the proportional importance of the input and output values normalised by the corresponding input and output values of the best unit observed. While this may sound a bit complicated, it is in fact very natural for the assessed unit to attach more weights to those inputs or outputs where its relative performance is high (even if there are many units performing better on the same dimensions).

As noted by Sarrico and Dyson (2004), it is easier to elicit from management virtual weights restrictions. The same is true for the weights in the models (22) and (23). Also, with respect to the model in (23), it is argued by Sarrico and Dyson (2004) that it makes more sense to impose proportional virtual weights restrictions only on the virtual inputs for an input-oriented model and only to the virtual outputs for an output-oriented model (this is essentially because the other set of weights do not add up to unity and is directly related to the efficiency score). This problem is not present in model (22) since both sets of weights add up to unity. Setting proportional virtual weights restrictions in classical DEA models is equivalent to setting simple restrictions on the weights in the models considered here. They would appear as $a_i \geq k_i$ (or $a_i \geq k_i$) and $b_j \geq l_j$ (or $b_j \geq l_j$) directly imposing lower or upper bounds to the proportional importance of the corresponding relative inputs and relative outputs. Due to the fact that both sets of weights in model (22) add up to 1, it is easy to convert any such simple restrictions into virtual assurance regions of type I, which are the most advocated forms of restrictions in Sarrico and Dyson (2004). The difference is only in appearance. For example, simple restriction $a_i \geq k_i$ can be easily converted into $a_1 + a_2 + \ldots + (1 - 1/k_i)a_i + \ldots + a_m \leq 0$. It is also possible to convert simple restrictions linking output weights and input weights into virtual assurance regions of type II. For example, $a_i + b_j \geq t$ can be converted into the following form: $a_1 + a_2 + \ldots + (1-1/t)a_i + \ldots + a_m - b_j/t \leq 0$ or $b_1 + b_2 + \ldots + (1-1/t)b_i + \ldots + b_s - a_i/t \leq 0$.

Let us now look at the variations of the basic models that allow for variable returns to scale. Clearly, *CCR* and *GE* are both constant returns to scale models. For *CCR* this is a well-know property and for *GE* this is obvious from its formulation in (17). Models (24) and (25) can be seen as input-oriented variable returns to scale variations of (22) and (23), respectively.

$$\begin{aligned}
&Max\ \theta_k, \\
&\text{s.t.} \\
&\sum\nolimits_j b_j \ln\left(\frac{y_{jp}}{y_{jk}}\right) - \sum\nolimits_i a_i \ln\left(\frac{x_{ip}}{x_{ik}}\right) + \theta_k \leq 0 \quad \forall p, \\
&\qquad\qquad \sum\nolimits_i a_i = 1, \quad a_i \geq 0 \\
&\qquad\qquad\qquad\qquad b_j \geq 0
\end{aligned} \tag{24}$$

$$\begin{aligned}
&Max\ \omega_k \\
&\text{s.t.} \\
&\sum\nolimits_j \beta_j \left(\frac{y_{jp}}{y_{jk}}\right) - \sum\nolimits_i a_i \left(\frac{x_{ip}}{x_{ik}}\right) + \beta_0 \leq 0 \quad \forall p. \\
&\qquad\qquad \sum\nolimits_i a_i = 1, \quad a_i \geq 0 \\
&\qquad \sum\nolimits_j \beta_j + \beta_0 = \omega_k, \quad \beta_j \geq 0
\end{aligned} \tag{25}$$

We can recall from the classical DEA theory that model (25) identifies increasing returns to scale for the unit assessed (k) if and only if $\beta_0 > 0$ for all optimal solutions and decreasing returns to scale if and only if $\beta_0 < 0$. These two conditions translate into $\Sigma\beta_j < \omega_k$ and $\Sigma\beta_j > \omega_k$, respectively. In a similar manner, the returns of scale in model (24) are increasing if and only if $\Sigma b_j < 1$ and decreasing if and only if $\Sigma b_j > 1$ for all optimal solutions. This can be also intuitively understood. For example, $\Sigma b_j < 1$ means that if all the current output levels of the assessed unit are multiplied by some scalar u then this will require multiplying all its input levels by $u^{(\Sigma bj)}$ to keep its efficiency score intact. But since $\Sigma b_j < 1$ then $u^{(\Sigma bj)} < u$, which means that the unit operates in conditions where an increase in outputs require less then proportionate increase in inputs, hence we have increasing returns to scale.

Relations of parameters a_i and b_j in model (22) and (24) to the main concepts from production theory, such as returns to scale, scale elasticities, rates of substitutions and marginal products are very interesting and important for using these models in practice. However, these issues will not be considered any further in this paper since they are all well covered in the existing literature on multiplicative models Banker et al. (2004), Banker and Maindiratta (1986). Banker and Maindiratta (1986) discuss production characteristics of models (22) and (24), which are presented in a slightly different form. Model (22) is referred to as the most productive scale size model in Banker and Maindiratta (1986) and is presented in its dual formulation. Model (24) is presented using output-orientation and with a slightly modified objective function. The forms used here are most suitable to understand their other advantageous properties such as their flexibility in modelling and their promising potential in ex-ante types of problems.

One of the most important properties of the geometric DEA is that it can easily deal with factors grouped into many categories, each of which may be structured into any number of hierarchical levels. To start with, it is relatively straightforward to visualise expansion of model (17) into any number of categories. Transforming such a model into a linear programming problem follows exactly the same steps as we used to obtain model (22). Weather a factor is of maximising or minimising nature should not play any role when grouping the factors into categories. We have already observed that *GE* model can treat inputs (normally minimising factors) as inverted outputs (normally maximising factors). Hence, it is plausible to invert all the minimising factors into maximising ones and then split the factors into categories based on the principle of substitutability. When faced with a decision making problem with many maximising and minimising criteria, we can collect similar factors into their own group irrespective of their maximising or minimising orientation. In this way, it is possible to put, for example, all environmental factors, all financial factor and all socially related factors into their own group. This is very natural since it makes much more sense to allow substitutability among the factors representing similar issues rather than to allow substitutability among the factors based on their measurement orientation (maximising or minimising). If in addition we want to alter the relative importance of any specific group, all what needs to be done is to alter the condition requiring that the sum of weights within each group is equal to 1. These sum-to-unity requirements for each group are equivalent to setting equal relative importance of each group in its contribution to the overall performance/desirability of the unit assessed.

The simplicity of the *GE* model and its weights are also the main reason why any criterion can be further split into a number of sub-criteria. To clearly see this, let us consider the hierarchical structure in Figure 1.

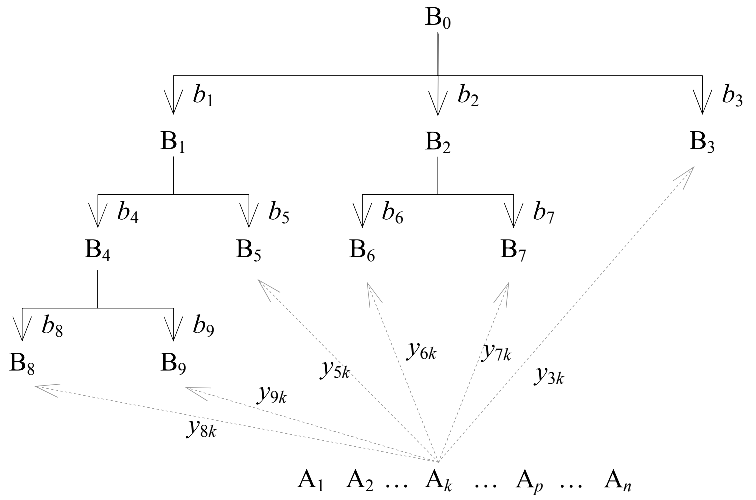

Figure 1. A hierarchy of criteria and sub-criteria with n units to be assessed (A_1, … A_n)

Without any loss in generality, we will assume that all the criteria are of the maximising type. Criteria are denoted as B_1, B_2, …, B_8 and their corresponding weights as b_1, b_2, …, b_9. B_0 is just the name of the overall goal or overall performance and does not require any weight. Units assessed are A_1, A_2, … A_n. Performances of all units are measured directly with respect to the criteria not being split further into sub-criteria. So, taking unit A_k, we have the following set of measures: y_{8k}, y_{9k}, y_{5k}, y_{6k}, y_{7k} and y_{3k}. Applying model (17) to calculate the overall performance of unit A_k, we have:

$$e_k = \max_{b_j \in B, b_j \geq 0} \min_p \left(\left(\left(\left(\frac{y_{8k}}{y_{8p}} \right)^{b_8} \left(\frac{y_{9k}}{y_{9p}} \right)^{b_9} \right)^{b_4} \left(\frac{y_{5k}}{y_{5p}} \right)^{b_5} \right)^{b_1} \left(\left(\frac{y_{6k}}{y_{6p}} \right)^{b_6} \left(\frac{y_{7k}}{y_{7p}} \right)^{b_7} \right)^{b_2} \left(\frac{y_{3k}}{y_{3p}} \right)^{b_3} \right). \tag{26}$$

Without any additional weights restrictions, the condition $b_j \in B$, used in the above formula represents normalisation of weights within each group: $b_8 + b_9 = 1$, $b_4 + b_5 = 1$, $b_6 + b_7 = 1$, and $b_1 + b_2 + b_3 = 1$. The expression in (26), however, can be simplified to the following form:

$$e_k = \max_{w_j \in W, w_j \geq 0} \min_p = \prod_{j \in \{\text{end criteria}\}} \left(\frac{y_{jk}}{y_{jp}} \right)^{w_j}, \tag{27}$$

where the weights w_j are the global weights of the end criteria. They are formed as the products between the corresponding b_j weights. In our example, $w_3 = b_3$, $w_5 = b_1 b_5$, $w_6 = b_6 b_2$, $w_7 = b_7 b_2$, $w_8 = b_8 b_4 b_1$ and $w_9 = b_9 b_4 b_1$. The w_j weights obtained in this way still have the same relative values when compared to the weights from its own group (the weights corresponding to the criteria belonging to the same parent criterion), so that we can take w_j weights to have the same meaning as the weights b_j for the end criteria. Normalisation of weights w_j follows directly from the normalisation of b_j weights within their own group. Now, any desired restrictions on weights b_j can be easily converted into the corresponding restrictions on w_j weights. For example, $b_4 \geq 2b_5$ translates into: $w_8 + w_9 \geq 2w_5$. So, expanding GE model to deal with any hierarchical structure, we are effectively getting a flexible multiplicative version of analytical hierarchy process (AHP) where the weights of criteria do not necessarily need to be specified in advance and can be specified through ranges if at all. Using more than two categories in classical DEA is treated mainly through a very specific problem where in addition to standard inputs and outputs we also have undesirable outputs (for detailed discussion on this issue see Thanassoulis et al., 2008). As for the treatment of factors in multiple levels, so far there was only treatment of the two-level DEA model (Meng et al., 2008; Kao, 2008).

Through this brief exposition of models and properties of the geometric DEA, it is the authors' hope that the flexibility and power of geometric DEA is made more apparent and that further research in the area as well as the use of these models in practice is well worth consideration.

REFERENCES

Banker, R.D., Cooper, W.W., Seiford, L.M, Thrall, R.M. and Zhu, J. (2004), "Returns to scale in different DEA models", *European Journal of Operational Research*, 154(2), pp. 345-362.

Banker, R.D. and Maindiratta, A. (1986), "Piecewise loglinear estimation of efficient production surfaces", *Management Science,* 32(1), pp. 126-135.

Charnes, A., Cooper, W.W. and Rhodes, E.L. (1978), "Measuring the efficiency of decision making units", *European Journal of Operations Research*, 2(6), pp. 429-444.

Despic, O., Despic, M. and Paradi, J.C. (2007), "DEA-R: ratio-based comparative efficiency model, its mathematical relation to DEA and its use in applications", *Journal of Productivity Analysis,* 28(1), pp. 33-44.

Kao, C. (2008), "A linear formulation of the two-level DEA model", *Omega*, 36(6), 958-962.

Meng, W., Zhang, D., Qi, L. and Liu, W. (2008), "Two-level DEA approaches in research evaluation", *Omega*, 36(6), pp. 950- 957.

Sarrico, C.S. and Dyson, R.G. (2004), "Restricting virtual weights in data envelopment analysis", *European Journal of Operational Research*, 159(1), pp. 17-34.

Thanassoulis, E., Portela, M. and Despic, O. (2008), "The mathematical programming approach to efficiency analysis", In H. Fried, K. Lovell & S. Schmidt (Eds.), *Measurement of Productive Efficiency and Productivity Growth* (pp. 251-420). Oxford: Oxford University Press.