Working papers series

WP ECON 07.09

Using a Modified DEA Model to Estimate the Importance Of Objectives. An Application to Agricultural Economics

Francisco J. André (U. Pablo de Olavide)
Inés Herrero (U. Pablo de Olavide)
Laura Riesgo (U. Pablo de Olavide)

JEL Classification numbers: C61, Q12
Keywords: Multicriteria Decision Making, Goal Programming, Weights, Preferences, Data Envelopment Analysis

Department of Economics

USING A MODIFIED DEA MODEL TO ESTIMATE THE IMPORTANCE OF OBJECTIVES. AN APPLICATION TO AGRICULTURAL ECONOMICS*

Francisco J. André **Inés Herrero** **Laura Riesgo****

Universidad Pablo de Olavide

ABSTRACT

This paper shows a connection between Data Envelopment Analysis (DEA) and the methodology proposed by Sumpsi *et al.* (1997) to estimate the weights of objectives for decision makers in a multiple attribute approach. This connection gives rise to a modified DEA model that allows to estimate not only efficiency measures but also preference weights by radially projecting each unit onto a linear combination of the elements of the payoff matrix (which is obtained by standard multicriteria methods). For users of Multiple Attribute Decision Analysis the basic contribution of this paper is a new interpretation of the methodology by Sumpsi *et al.* (1997) in terms of efficiency. We also propose a modified procedure to calculate an efficient payoff matrix and a procedure to estimate weights through a radial projection rather than a distance minimization. For DEA users, we provide a modified DEA procedure to calculate preference weights and efficiency measures which does not depend on any observations in the dataset. This methodology has been applied to an agricultural case study in Spain.

Keywords: Multicriteria Decision Making, Goal Programming, Weights, Preferences, Data Envelopment Analysis.

---

* We thank E. Thanassoulis, Victor Podinovski and Diego Prior for useful comments. We also gratefully acknowledge financial support from the European Commission (research project EFIMAS, Proposal no. 502516) and the Spanish Ministry of Education and Science (research project SEJ2006-08416/ECON) F.J. André also thanks project SEJ2005-0508/ECON.

** Corresponding author. Tel: +34 954 349 851; Fax: +34 954 349 339. E-mail: laurariesgo@upo.es. Address: Dept. of Economics. Pablo de Olavide University. Ctra. de Utrera, km. 1 - 41013 Sevilla. Spain.

1. Introduction and objectives

Several authors have pointed out some close connections between Data Envelopment Analysis (DEA) and Multicriteria Decision Making (MCDM) See Belton and Vickers (1993), Stewart (1994, 1996), Joro *et al.* (1998), Bouyssou (1999). These authors have underlined the equivalence between the notion of 'efficiency' in DEA and MCDM (see, for example, Bouyssou, 1999, p. 974) although both approaches are different regarding how efficiency is measured in practice. In DEA, the so-called 'efficient frontier' is built as the envelope of all the Decision Making Units (DMUs hereafter) included in the sample, so that efficiency is measured in relative terms by comparing each unit with the rest in the sample. On the contrary, efficiency is measured in absolute terms in MCDM. That is, in a MCDM problem, the decision maker (DM) faces a number of constraints which determine the feasible set. Therefore, by exploring the feasible set it is possible to determine which solutions are efficient or not (and hence, which DMs taking those solutions behave in an efficient way, without any comparison across DMs). In order to translate multicriteria objectives into DEA terminology, a "*max*" objective can be understood as an output whereas a "*min*" objective can be interpreted as an input or a bad output (see Doyle and Green, 1993; Steward, 1994 and Bouyssou, 1999).

Further to the efficiency concept, in this paper we report an additional connection by stressing the parallelism between DEA and the multicriteria methodology proposed by Sumpsi *et al.* (1997) to estimate the weights of different objectives in the preferences of DMs. We claim that, although these methodologies have been developed independently from each other, there is a strong parallelism between them.

MCDM and DEA also have in common that both of them deal with individuals, activities or organizations that are concerned with multiple objectives or inputs and outputs. In such a framework, it seems relevant to measure or to evaluate the relative importance of each objective, input or output according to the preferences of DMs. The methodology by Sumpsi *et al.* is aimed at measuring this importance by projecting the observed values of objectives onto a linear combination of the elements of the payoff matrix (where such a matrix is obtained by optimizing each objective separately). Since these elements are efficient (see below for further discussion), we claim that this procedure has a strong resemblance to DEA, where each unit is projected on a combination of efficient units. The first contribution of this paper is to underline this connection, hence providing a new interpretation for the Sumpsi *et al.* procedure.

On the other hand, although the aim of DEA is not estimating preferences but efficiency scores, it requires constructing a weighted combination of inputs and outputs. As the weights (known as *virtual multipliers*) used to compute such combinations are endogenously determined to provide the best possible score for each unit, they could be understood as having some connection with the preferences of DMs. For example, Cooper *et al.* (2000, section 6.6) suggests to bound DEA weights according to the importance given by some experts to each of the criteria (inputs) using an Analytic Hierarchy Process (AHP) analysis. Nevertheless, the weights obtained from a standard DEA analysis do not represent a suitable measure of the preferences of a given DM, since DEA parameters are crucially influenced by the structure of the production process under analysis, which often is just related to technological issues, and the representation of the efficient frontier is crucially influenced by the amounts of inputs and outputs of other observations in the data.

In this paper we try to establish a particular way to apply DEA in order to obtain estimates of preference parameters, by taking advantage of the parallelism between DEA and the Sumpsi *et al.* methodology. For this purpose, we propose to project each decision unit onto a linear combination of the elements of the payoff matrix. The rationale behind this procedure is to control for the technological constraints (those related to the production structure) and isolate the effects specifically associated to preferences. The payoff matrix allows constructing a linear version of the set of efficient alternatives among which it is possible to choose. By evaluating the distance to each element of the payoff matrix it can be inferred which criteria are revealed as preferred for the DM. Using this approach, we get both an estimation of the preference weights for each DM and a measure of efficiency in a single model. This efficiency measure has the property of being independent of the rest of the observations in the dataset.

The paper has the following structure: Section 2 reviews the basic elements of the DEA approach. Section 3 presents the Sumpsi *et al.* methodology and proposes a modification to guarantee that all the elements of the payoff matrix are efficient. The fourth section offers an alternative way to use DEA in order to measure efficiency and estimate the weights of inputs and outputs. Section 5 presents an empirical application of the suggested method to agricultural economics using real data from an irrigated area in Spain. We obtain efficiency measures that are very close to the real values and to conventional DEA measures. We also obtain preference weights that are very similar to those obtained when using the Sumpsi *et al.* methodology. To test the practical usefulness of these estimates we show, in a validation exercise, that they provide a good approximation to observed behaviour. Section 6 summarizes the main contributions of the paper.

2. DEA model

In a standard DEA model there are n DMUs, using s different inputs to produce t different outputs. Using the standard notation, X_{ij} and Y_{rj} denote, respectively, the amount of input i used and output r produced by the $j-th$ DMU. X_j (Y_j) represents the $s\times 1$ ($t\times 1$) column input (output) vector corresponding to the $j-th$ DMU. The DEA model proposed by Banker *et al.* (1984), known as BCC model after its authors, allows for variable returns to scale, by forcing the weighting parameters $\lambda^T \equiv (\lambda_1,\ldots,\lambda_t)$ to add up to one (see below) in the original model of Charnes *et al.* (1978). In order to measure the efficiency of a specific DMU, labelled as '0', the following output oriented (dual) model has to be solved:

$$\begin{aligned}
&Max \quad \theta \\
&s.t.: \\
&\lambda^T\, Y \geq \theta Y_0 \\
&\lambda^T X \leq X_0 \\
&\vec{1}\lambda = 1 \\
&\lambda \geq 0
\end{aligned} \qquad \text{(BCC}_\text{D}\text{-O)}$$

where X (Y) is the matrix representing all the inputs (outputs) of all the DMUs, T denotes transposing, $\vec{1} \equiv (1,...,1)$ and the λ_j parameters ($j=1,\ldots,n$) are the weights associated to each observed DMU in order to construct a convex combination of all of them (or just a subset if some λ_j's are equal to zero). The values of these parameters are DMU-specific.

DEA seeks to identify efficient units and combine them to construct an *efficient frontier*. A unit is said to be *radially* efficient if the optimal value of θ is equal to one. In order to guarantee that a unit is fully efficient, a second phase analysis should be carried out. In this second optimisation stage the sum of the positive and negative slacks, defined as $s^+ = \lambda^T Y - \theta Y_0$ and $s^- = X_0 - \lambda^T X$ respectively, is maximised. In this case, a unit is said to be *fully* efficient if the optimal value of θ is equal to one and all the slacks are equal to zero. The *peer units* associated to the unit under analysis are those with a strictly positive value of λ. The combination (weighted by the λs) of these peer units defines a virtual unit on the frontier. We could make the unit under analysis be efficient by transforming it into this virtual one.

We can also interpret DEA as minimizing the distance from the unit under analysis to the set of hyperplanes that envelopes all DMUs. This interpretation is more easily understood using the (output oriented) primal model, which has the following formulation:

$$\begin{aligned}
&Min \quad v^T X_0 - v_0 \\
&s.t.: \\
&\mu^T Y_0 = \vec{1} \\
&-\mu^T Y_j + v^T X_j - v_0 \vec{1} \geq 0 \quad \forall j = 1,\dots,n \\
&\mu, v \geq \vec{0}
\end{aligned} \qquad (\text{BCC}_\text{P}\text{-O})$$

where coefficients μ_i and v_r are known as *virtual multipliers*.

Inefficient DMUs can be projected onto the efficient frontier by radially expanding their outputs or radially contracting their inputs for an output or input orientation respectively. In Figure 1 an example has been graphed for an output-oriented model using one input (the amount of which is assumed to be the same for all units) to produce two outputs. Inefficient unit E can be projected onto point E' on the efficient frontier and the same occurs for F and F'. The expansion factor for point E is defined as the ratio of distances $\theta = OE'/OE$. This value is lower bounded by one and a value equal to one means that the virtual unit on the frontier is the evaluated unit itself, and hence it is efficient. The technical efficiency rate (TE) is given by $TE = 1/\theta$, which is upper bounded by one and lower bounded by zero.

Andersen and Petersen (1993) proposed a modified version of the original BCC model, commonly used to rank efficient units, where the DMU under evaluation is excluded from the reference set, so that it may be above the frontier. Inefficient units get the same score as in the standard model, whereas efficient ones get a so-called *super-efficiency* score that can be larger than one. The larger it is, the more efficient the associated DMU is. Units with an efficiency score greater than one are said to be *super-efficient*.

Since the virtual multipliers μ_i and v_r from (BCC$_P$-O) are endogenously determined to provide the best score for each DMU, they could be interpreted as being somehow related to the *weight*, or the *importance*, the unit under analysis gives to each input or output in order to achieve the maximum efficiency. See Cooper *et al.* (2000), (page 25 and pages 169-173), and Tone (1989, 1999). Nevertheless, these coefficients cannot be interpreted as measuring the preferences of DMs, since they are basically technical parameters (see Allen *et al.* 1997). Below we show an alternative way to use DEA in order to estimate the importance of each input and output for each DMU. The idea is to take advantage of the parallelism between DEA and the methodology by Sumpsi *et al.*, which is summarized in the following section.

3. Estimating the weights of attributes in a multiple attribute context

The methodology proposed by Sumpsi *et al.* (1997), and extended by Amador *et al.* (1998) is based upon weighted goal programming and allows estimating the weight or the importance that different objectives have on the observed behaviour of DMs to be compatible "not with the answers (…) to artificial questionnaires but compatible to the actual behaviour which they follow" (Sumpsi *et al.* 1997, page 65).

Assume that some DM have a set of q "more is better" objectives[a] depending on a vector of decision variables x according to the functions $f_i(x)$ ($i = 1,\ldots,q$). Moreover, the DM faces a number of known technical constraints which somehow limit his decisions. The first step is to construct the payoff matrix for these objectives. The first element of the first column of this matrix is obtained by solving the problem:

$$\begin{aligned} &Max \quad f_1(x) \\ &s.t.: \\ &\quad x \in F \end{aligned} \qquad [1]$$

where F denotes the feasible set which results from problem-specific constraints. The optimal value of $f_1(x)$, denoted as $f_1^* \equiv f_{11}$, is the first entry of the payoff matrix. To get the other entries of the first column, we substitute $arg\,max\, f_1(x)$ in $f_i(x)$, for $i = 2,\ldots,q$. The rest of the columns of the payoff matrix are obtained by implementing

[a] Note that this assumption does not imply any loss of generality. A "less is better" objective can be transformed in "more is better" multiplying by –1. If the target is to get exactly a certain value, the objective can be written as minimizing the distance (or maximizing the opposite of the distance) from the attained value to the target value, so that it can be formulated as a "less is better" (or "more is better") objective. Therefore, this formulation permits us to deal with any problem involving any of the relevant types of objectives.

the same kind of calculations, i.e., the generic element f_{ij} is obtained by plugging $arg\,max\ f_j(x)$ in $f_i(x)$.

Notice that the payoff matrix may not be unique since [1] could have alternative optima, and some of them could be inefficient. To illustrate this, assume there are two objectives and the feasible set is represented by the polygon OABCD in Figure 2. When optimizing objective 1 (2), we could obtain any point on segment CD (AB). Since we are interested in an interpretation in terms of efficiency, it is convenient to have efficient points as a reference. In Figure 2, the set of efficient solutions is given by segment BC, so we should select point C for the first column of the payoff matrix and B for the second column. We propose to do this by solving the following lexicographic problem for every objective $i = 1,\ldots,q$:

$$\begin{aligned} &Lex \quad max\left\{ f_i(x), \sum_{j\neq i} \alpha_j f_j(x) \right\} \\ &s.t.: \\ &\qquad x \in F \end{aligned} \qquad [2]$$

meaning that objective i is maximized and, if some alternative optima exist, an arbitrary linear combination of the rest of objectives (with $\alpha_j > 0$ for all $j \neq i$) is optimized without worsening the performance of objective i. For our purpose, the specific values of α_j do not matter, as long as they are positive and, henceforth, they provide an efficient solution. In our application we calculate the payoff matrix using this procedure. Hence, from now on, we will assume that all the columns of the payoff are efficient by construction.

Assume now that the researcher observes the decision made by a specific DM facing the decision problem described above. Using this information, the researcher aims at eliciting the weights given by the DM to each objective. Following Sumpsi *et al.* (1997) this elicitation can be done by solving the following system of $q+1$ equations:

$$\begin{aligned} &\sum_{j=1}^{q} w_j f_{ij} = f_i \qquad i = 1,\ldots,q \\ &\sum_{j=1}^{q} w_j = 1 \end{aligned} \qquad [3]$$

where f_i is the observed value of the $i-th$ criterion and w_j measures the weight of the $j-th$ objective. Usually, a positive solution to system [3] does not exist, so that it is necessary to find the closest set of weights by solving the problem:

$$
\begin{aligned}
& Min \sum_{i=1}^{q} \frac{n_i + p_i}{f_i} \\
& s.t.: \\
& \sum_{j=1}^{q} w_j f_{ij} + n_i - p_i = f_i \qquad i = 1,\ldots,q \\
& w_j \geq 0 \qquad\qquad\qquad i = 1,\ldots,q \\
& n_i, p_i \geq 0 \qquad\qquad\quad i = 1,\ldots,q \\
& \sum_{j=1}^{q} w_j = 1
\end{aligned}
\qquad [4]
$$

where n_i (p_i) is the negative (positive) deviation variable from the observed (real) value f_i. When solving [4], the observed point is projected onto another point that is constructed as a weighted sum of the elements of the payoff matrix.

The first key insight of this paper is the strong parallelism of this methodology with DEA once it is guaranteed that the elements of the payoff matrix are efficient. Given that, by construction, the elements of the payoff matrix are efficient, *the solution of* [4] can be interpreted as projecting every observation onto a combination of efficient units. The main difference with respect to DEA is that the reference units are not “real” observed units, but potential (feasible) observations that could show up if the DM were interested in maximizing just one objective. A second difference is that, by construction, in this procedure, the approximation to the efficient frontier is linear (instead of piece-wise linear as it is usual in DEA).

Figure 3 illustrates a problem with two objectives where the feasible set is defined by the polygon ABCDE. Problem [4] consists of finding a point on segment *AB* as close as possible to the observed vector x. Point x could be projected on F or G depending on the slope of segment *AB* and on the observed value of f_1 and f_2. This is a consequence of using L_1 (“Manhattan”) metric in [4]. Different metrics result in different projections. For example, the L_2 or Euclidean metric (implemented by minimizing a combination of the squared deviation variables, instead of the variables themselves) would result in an orthogonal projection of x on segment *AB*.

Another important insight is the fact that *the* w_j parameters resulting from problem [4] can be understood as representing the weight of each criterion in the preferences of the DM. The interpretation is the following: if an agent faces the decision problem depicted in Figure 3, he can choose among all the feasible points in ABCDE. If he chooses one specific point and discards all the rest, he is *revealing* which alternative he prefers. This follows a standard reasoning based on the Revealed Preference Theory (see, for example, Mas-Colell *et al.* 1995, chapter 1).

In this case, we observe that the DM chooses point x, which seems to indicate that this point represents his most preferred option. Nevertheless, point x itself cannot be understood as being the result of a rational decision making process since it is inefficient and we know that "Paretian efficiency is a necessary condition to guarantee the rationality of any solution provided by any MCDM approach" (Ballestero and Romero 1998, p. 7). Once x is projected onto (our approximation of) the efficient frontier, the resulting projection (in this case, point F or G) can be taken as a surrogate of the observed decision, i.e, that efficient point which is as close as possible to the observed one. The distance from x to its projection can be interpreted as inefficiency of the DM and the projected point is, by construction, a convex linear combination of A and B. If the DM is concerned only about the first (second) objective, he should choose point B (A) or a point very close to it, so that we should obtain $w_1 = 1$, $w_2 = 0$ ($w_1 = 0$, $w_2 = 1$). Similarly, if the first (second) objective is more important than the second (first) one, the observed decision should be closer to B (A) than A (B). In general, when objective j is very important (is not very important) for the DM, the observed vector of achieved objectives should be very close (not be very close) to f_j and therefore w_j should be very close to 1 (to 0). Using this method we are measuring *revealed preferences*, as opposed to *declared preferences*, which are typically obtained from direct surveys. For an application of this method to estimate preferences see, for example, Gómez-Limón and Berbel (2000) or Gómez-Limón and Riesgo (2004).

4. Combining methodologies: using a modified DEA model to estimate the weights associated to each throughput

Using the parallelism between the methodologies presented in section 2 and 3, our aim is to find a way to use DEA in such a way that it provides a measure of preference parameters. Although DEA virtual multipliers μ_i and v_r from problem (BCC$_P$-O) are associated to outputs and inputs respectively, they cannot be properly interpreted as preference parameters linked to these outputs and inputs, because they are affected by the technological structure of the activity under analysis and by the values of these throughputs for the other units. To get this point, assume a case with two outputs (Y_1, Y_2) and consider a specific DMU focused on maximizing only Y_1 and not caring at all about Y_2. These preferences should be represented by a weight equal to one for Y_1 and zero for Y_2 ($w_1 = 1$, $w_2 = 0$). Nevertheless, it may well be the case that the feasible set is such that the minimum attainable value of Y_2 is strictly positive. As a consequence, we could observe that this DMU has a positive value for Y_2 and we may obtain a strictly positive value for the virtual multiplier associated to Y_2. However, this positive value should not be interpreted as a positive preference for output 2, as it is determined by technical issues, i.e., by the shape of the feasible set.

Furthermore, in DEA, efficiency is measured in relative terms, in the sense that the efficiency score depends on the observations the unit under analysis is compared to, and the values of the virtual multipliers also depend on the reference set. Nevertheless, the preferences of a DM, as they are typically understood in economics or in decision theory, are privately given and do not depend on the rest of agents.

In order to get a measure of preference parameters, we suggest using a modified BCC$_D$-O DEA model with the only difference that the reference set is not made up of all observations in the data set, but instead is made up of the elements of the payoff matrix, i.e., those extreme (virtual) units optimizing every criterion separately. We will call it modified DEA model.

Therefore, the model would take the form:

$$\begin{aligned}
&Max \quad \theta \\
&s.t.: \\
&\lambda^T Y^* \geq \theta Y_0 \\
&\lambda^T X^* \leq X_0 \\
&\vec{1}\lambda = 1 \\
&\lambda \geq 0
\end{aligned} \qquad [5]$$

where Y^* is a matrix which rows contain the value of the outputs (values of the objectives for maximising) for each of the elements in the payoff matrix. Similarly X^* is a matrix where in each row there are the values of the inputs (or criteria to be minimised) of a given unit of the payoff matrix. Therefore Y^* and X^* have the same amount of rows as the number of elements of the payoff matrix. The rest of the elements in problem [5] are the usual in a standard DEA model.

By using this modified DEA model, the values of λ associated to each unit of the payoff matrix have a particular meaning: they can be considered as estimates of the preference weights given to each objective (input/output). Note that we are projecting each observation on a convex combination of the elements of the payoff matrix, so that the values of λ represent the degree of proximity of the observed unit to each of these virtual units associated to the maximization (or minimisation) of each of the different criteria.

The rationale behind this procedure is the following: the elements of the payoff matrix explicitly recognize that, when optimizing just one objective (or equivalently, one output or input), each DMU may have to take a certain value of the rest of attributes for technological or feasibility reasons. When including these elements in the reference set, the resulting coefficients represent the importance that the unit under analysis gives to each of the criteria controlling for the feasibility constraints. Furthermore, as the reference elements are efficient by construction, the hyperplane connecting them can be taken as an approximation of the efficient frontier, and the distance from each DMU turns out to be an alternative efficiency measure with the property of being independent of any DMUs in the sample.

Concerning the selection of the BCC_D-O model, choosing an adequate version of DEA is not a trivial task. In this case, we aim at stressing the parallelism of DEA with the Sumpsi *et al.* methodology. There are at least two types of models that may be applied: additive models (Charnes *et al.*, 1985 and Tone, 2001) or conventional radial models. For consistency, the former should be compared to the Sumpsi *et al.* model using a L_1 norm while the latter should be compared to the Sumpsi *et al.* model using a L_2 norm, given that a radial expansion to the frontier is generally closer to a L_2 norm than to a L_1 norm. In the application presented below, we preferred to use an output oriented radial model to an additive model because, for comparison purposes, additive models present the disadvantage that they *maximize* the slack variables (i.e., they *maximize* the L_1 distance to the frontier, instead of *minimizing* this distance, as in the Sumpsi *et al.* methodology). For this reason we preferred the use of DEA radial models as they minimize the radial expansion (or contraction) to the efficient frontier.

Specifically, in our application (see section 5) we have used a BCC output-oriented DEA model because, as discussed below, in the preferences of most DMs, profit maximization seems to be the key element and this appears to match with an output-oriented approach. Furthermore, we consider super efficiency (i.e., we do not include the unit under analysis in the reference set) to guarantee that the projection of any point (originally below or above the frontier) is always a combination of the elements of the payoff matrix.

Comparing methodologies

Figure 4 compares the results from standard DEA, the Sumpsi *et al.* methodology and the modified DEA model proposed above. In standard DEA, the reference set contains all the observed DMUs (represented by black dots). The efficient frontier is constructed as the envelope of all these units (in Figure 4, FDJEI) and the efficiency of each unit is measured as the distance from it to the frontier when radially projected. In the Sumpsi *et al.* methodology the reference set consists only of the elements of the payoff matrix, which in Figure 4 correspond to points A and B (marked with a star) and the goal is to find a linear convex combination of these elements as close as possible to the observed units according to some metric (in the figure, we illustrate the L_2 metric). We propose a combination of both methods by taking the payoff matrix as the reference set and projecting each unit radially on it. For example, unit C is projected on point C'' when using the Sumpsi *et al.* methodology and it is projected on point C' when using the modified DEA method (which, in this specific case, by coincidence, equals the standard DEA projection). A similar exercise is made for point E. Since E is efficient, it is projected on itself when using DEA, on E'' when using the method developed by Sumpsi *et al.* and on E' when using the modified DEA method.

Compare, first, the results for modified DEA and Sumpsi *et al.* methods. In some cases, such as point D, both projections are virtually the same but in other cases (e.g. point E) there are some differences due to the different projection criteria used in both approaches: in the case of Sumpsi *et al.* it consists of minimizing the distance whereas, by following DEA the projection aims at keeping the proportions of outputs unchanged. Obviously, the size of the difference depends on the specific example, but *a priori* they should not be very different for most cases and, specifically, they should be very similar for “average units” with a balanced combination of outputs (i.e., those units not very far from the 45-degrees line). In fact, the application in the case study shows very similar preference parameters from both approaches.

Using the modified DEA approach, we also obtain an efficiency measure as the distance from each unit to the new frontier AB, so that we can compare this measure to standard DEA. For example, the efficiency score for point C is the same in a standard DEA approach and in the modified DEA method (being inefficient in both cases). Units D and E, which appear to be efficient in a standard DEA method, appear to be super-efficient in the modified DEA method. Nevertheless, in the application presented below it is shown that, although the numerical value of the efficiency scores can be different for standard and modified DEA model, the rankings of units tend to be rather similar

Table 1 presents a summary of the key features of the three methods. The second, third and fourth columns display the information requirements for each method. All three procedures require information about inputs and outputs (objectives in MCDM terminology) for the DMU under analysis (DM in MCDM terminology). In standard DEA this information is required not only for the DMU under analysis, but for all the DMUs of the sample. Concerning information requirements, the modified DEA method is equivalent to the one by Sumpsi *et al.*. in the sense that it does not require any sample but does require the payoff matrix, which in turn typically requires information about the structure of the decision problem, i.e., the relevant objectives and the constraints faced by the DM.

The fifth and sixth columns display the information provided as an output by each method. In this respect, DEA is basically aimed at providing just efficiency measures whereas the Sumpsi *et al.* methodology provides just preference weighting parameters. In this respect, the modified DEA approach amounts to a combination of both methods but giving both pieces of information.

Finally, the last column underlines the criterion that is used to project each unit on the frontier. In the Sumpsi *et al.* method the projection is done by minimizing the distance from the observed point to its projection. The modified DEA method follows the usual DEA spirit by using a radial projection.

5. Application to agricultural economics and a case study

5.1. Framework

A number of authors have pointed out that, contrary to the usual assumption in conventional economics, farmers are not only concerned with the maximization of profit, but other attributes such as risk, management complexity, leisure time, indebtedness, etc., are also involved in farmers' decision making. See Gasson (1973), Smith and Capstick (1976) or Cary and Holmes (1982). More recently, Willock *et al.* (1999), Solano *et al.* (2001) have also stressed this point.

Since farmers make their decisions trying to simultaneously optimize a range of conflicting objectives, we analyzed the behaviour of farmers under the MCDM paradigm. Specifically, we used the theoretical framework of multiatribute utility theory (MAUT). As pointed out by Herath (1981) or Hardaker *et al.* (1997, p. 162), the main drawback of this approach comes from the elicitation of the multi-attribute utility function (MAUF), including the mathematical shape of utility functions and the estimation of the weights of each attribute. Concerning the former issue, we assume an additive and linear MAUF. For a justification of this assumption, as well as its limitations, Gómez-Limón *et al.* (2003) and Gómez-Limón and Riesgo (2004) can be consulted. The resulting expression for the MAUF is

$$U = \sum_{j=1}^{q} \frac{w_j}{k_j} F_j(x) \qquad [6]$$

where U is the utility obtained by the DM, F_j is the value of attribute j, k_j is a normalizing factor (usually the observed value of each attribute *j*), w_j is the weight of attribute j, and x is the vector of decision variables.

Weights for different objectives are widely used in MCDM but there is some vagueness about exactly how these weights should be interpreted. Using the MAUT approach gives us a precise interpretation for these weights as the marginal utility of each (normalized) attribute. More details about the MAUT approach can be found in Keeney and Raiffa (1976), Edwards (1977), Farmer (1987), Amador *et al.* (1998), Ballestero and Romero (1998) or Huirne and Hardaker (1998).

Concerning the estimation procedure, we are interested in comparing the Sumpsi *et al*. methodology -which has been successfully checked in a number of studies, such as Berbel and Rodríguez (1998), Arriaza *et al.* (2002), or Gómez-Limón and Riesgo (2004)- with the modified DEA approach suggested above.

5.2. Case study and data set

The case study is a sample of 61 farmers from the community of irrigators Canal General del Páramo, located in northern Spain. This area has 15,554 irrigated hectares (ha), divided among 5,950 landowners. It has a "mild Mediterranean" climate, 800 m above sea level, with long, cold winters and hot, dry summers. Rain falls mostly in spring and autumn. In decreasing order of importance, the normal crop mix is maize, winter cereals, beans and set-aside. All the data to feed the models were obtained both from official statistics and from a survey developed in the area under study during the 2000-01 agricultural year. For more information about the survey see Gómez-Limón and Riesgo (2004). To simulate the decision-making process of farmers under the MAUT framework, we construct a mathematical model where farmers decide the value of some decision variables, being limited by certain constraints, in order to optimize various objectives:

Decision variables. Each farmer has a vector $x \equiv (x_1,\ldots,x_4)^T$ of decision variables determining the crop distribution. Variable x_h ($h = 1,\ldots,4$) measures the amount of land devoted to every particular crop, h, including winter cereals, maize, beans and set-aside. To get a normalized solution, we assumed that total land size of a farmer is 100 ha.

Constraints. We identify the following constraints as applied to each farmer:

- *Land constraint*. The sum of all crops must be equal to the total surface available to each farmer, which is normalized to 100 *ha*:

$$\sum_{h=1}^{4} x_h = 100 \qquad [7]$$

- *Common Agricultural Policy* (*CAP*) *constraints*. To fulfil the CAP requirements, we included 20% of set-aside for cereal, oilseed and protein crops. Any land devoted to set-aside greater than this percentage is excluded from EU subsidies, and this is taken as an invalid option in the model:

Maximum set aside: $x_4 \leq 20\% \cdot (x_1 + x_2)$ [8]

- *Rotational constraints*. These were taken into account according to the criteria revealed by the farmers in the survey. For rotational conditions, farmers do not usually crop winter cereals during two consecutive years in the same land. To represent this constraint we assume that the maximum area devoted to winter cereals in a year is half the total surface available:

$$x_1 \leq 50\% \cdot 100 \qquad [9]$$

Objectives. After the survey developed in the area under study, we concluded that farmers take the following objectives into account:

- *Maximization of total gross margin* (*TGM*), as a proxy of profit since, in the short run, the availability of structural productive factors (land, machinery, etc.) cannot be changed and financial viability of farms basically depends on gross margin. TGM data are obtained from the average crop margins in a time series of seven years (1993/1994 to 1999/2000) in constant 2000 euros:

$$TGM = \sum_h GM_h \cdot x_h \qquad [10]$$

where GM_h is a technical coefficient measuring the gross margin per unit of crop *h*.

- *Minimization of risk* (*VAR*). As noted by several authors, (see for example Just, 1974, Young, 1979, and Gómez-Limón *et al.* 2003), farmers typically have a marked aversion to risk, so that risk is an important factor in agricultural activity. Following the classical Markowitz (1952) approach, risk is measured by the variance of TGM: $VAR = x^T \left[COV\right] x$, where $\left[COV\right]$ is the variance-covariance matrix of the crop gross margins obtained from different crops, during the seven-year period. This classical approach has also been used in some recent works as Francisco and Ali (2006), Gómez-Limón and Martínez (2006) or Bazzani (2005).
- *Minimization of total labour input* (TL). This objective implies not only a cost reduction, but also an increase in leisure time and the reduction of managerial involvement, since labour-intensive crops require more technical supervision. Total labor requirements are calculated in the following way:

$$TL = \sum_h L_h \cdot x_h \qquad [11]$$

where L_h represents the technical coefficient representing the labour needs (hours per hectare) for each crop h.

To translate these objectives into DEA terminology, note that a "*max*" objective can be understood as an output (with the exception of "bad outputs") whereas a "*min*" objective can be interpreted as an input or a bad output. There are several ways to deal with bad or undesirable outputs (see for example Scheel, 2001). In this application, we use an output-oriented DEA model where the criterion to be maximized (gross margin) is considered the only output and the criteria to be minimized are treated as inputs (see Doyle and Green, 1993, Steward, 1994, and Bouyssou, 1999).

Using observed values of the crop distribution for every farmer, and the relevant technical coefficients (see equations [10] and [11]), we can compute the expected values for the objectives. Moreover, we introduced an artificial inefficiency component in the data so that we can test the ability of the model to measure efficiency by comparing the real (artificially introduced) efficiency rate with the estimated efficiency. We randomly generated 61 values ξ_{i} (i=1,…,61) from a normal distribution with mean 0.95 and standard error 0.10, and we multiplied the TGM of each farmer by the truncated version $\overline{\xi}_i = min\{\xi_i, 0\}$, so that we associated to each observation an efficiency score equal to the resulting (truncated) random number, with an average efficiency 0.913 and standard error 0.085.

5.3. Results

First, the estimated preference parameters using both the Sumpsi *et al.* methodology (with Euclidean metric) and the modified DEA approach are compared. Using the Sumpsi *et al*. approach, total gross margin (TGM) turns out to have a weight, w_1, greater than 0.5 for approximately 82% of the farmers, while $w_1 > 0.9$ for some 12% of them. For risk (VAR), the percentages are 18% and 0%, respectively. Total labour (TL) appears as a relevant objective for only 16% of the sample.

When estimating the weights (λ) with the modified DEA method, we also obtain that TGM is the most important objective ($\lambda_1 > 0.5$) for 82% of farmers, whereas for 27% of the sample, the weight of this objective is $\lambda_1 > 0.9$. In the case of VAR, we observe that

18% of farmers attach a weight greater than 0.5. Finally, with respect to TL, none of farmers seem to regard total labour minimization as a relevant objective.

Figures 5 and 6 show the cumulative distribution function of weights and Table 2 shows some descriptive statistics. We can see that the results from both approaches are very close. The correlation coefficient between weights using both methodologies is 98.5% for TGM and 97.6% for VAR. With regard to TL, the weights are zero or very close to zero for most of the farmers with any of the methods. Table 2 also shows the average differences between the weights calculated by both methodologies: 0.05 for TGM, 0.03 for VAR and 0.03 for TL. We conclude that the elicitation of farmers' preferences using Sumpsi *et al.* or the DEA modified version is virtually identical in this exercise.

To test the accuracy of these estimates, the following validation exercise is performed: substituting the estimated weights (we use those obtained from modified DEA model although the results are virtually the same when using Sumpsi *et al*.) and the mathematical expressions of the attributes in [6], we simulated the behaviour of farmers by maximizing farmers' utility subject to the constraints. Then, we compared the simulated values of both the decision variables and the objectives to those in the real observed situation, as it is usually done in validation exercises (see, for example, Qureshi *et al*., 1999). As shown in Table 3, the deviation between the average values for the objectives and the decision variables is small enough to permit us to regard the estimated model as a good approximation to the actual decision-making process.

Table 4 displays the results on efficiency measures. The modified DEA model provides a set of efficiency scores with mean 0.95 and standard error 0.075, so the DMUs appear to be slightly more efficient with our method as compared to the artificial inefficiency values and to standard DEA scores. This small difference can be understood as the effect of using a linear approximation to the efficient frontier. Nevertheless, the scores from modified DEA model turn out to be highly correlated (0.83) with the real inefficiency values and to those generated with standard DEA (0.83), so they seem to provide an acceptable inefficiency measure with the additional advantage of being independent on the set of DMUs in the sample.

6. Conclusions and further research

This paper reports a further link between DEA and MCDM, apart from those previously reported in the literature. Specifically, we have pointed out the parallelism between DEA and the MCDM methodology proposed by Sumpsi *et al*. (1997) to estimate the weights of different objectives for the DMs. Firstly, we have shown that the Sumpsi *et al*. method has a close connection with DEA in the sense that it can be seen as projecting every observation on a linear combination of efficient units. To guarantee that this is an accurate statement, we have proposed to construct the payoff matrix by solving an auxiliary lexicographic problem.

This connection can be exploited in order to suggest a modified version of DEA in order to measure preference weights. The main idea is to use DEA including the elements of the payoff matrix as the only units in the reference set and interpret the λ parameters as the weights of each criterion or throughput. The purpose of this technique is to account for the effect of technological (feasibility) constraints in the decision making process. This way a single technique is capable of providing estimates of preference parameters and an alternative efficiency measure with the property of being independent of the DMUs in the sample.

We have illustrated the suggested approach with an application to agricultural economics. The results show that the weights provided by the Sumpsi *et al*. methodology and the modified DEA model appear to be virtually identical and to provide a good approximation to actual decision making of the individuals in the sample. Moreover, the inefficiency measures provided by the modified DEA method turn out to be very close to the real values artificially introduced in the data, and also very close to the results obtained from a standard DEA approach.

Taking into account the information in Table 1, we can clarify the practical contribution of the method for each user. For MCDM users we have shown a new way to understand the method suggested by Sumpsi *et al.* (1997) in terms of efficiency: the projected point can be seen as a combination of efficient units. Moreover, we have proposed a modified procedure to calculate the payoff matrix to guarantee that all its elements are efficient. Finally, we propose to estimate the weights by making a radial projection rather than minimizing the distance. This procedure has the property of keeping the objectives ratio unchanged, which, in some situations, could provide a better approximation for the *true* preferences. For DEA users, we have provided a modified DEA procedure which allows calculating weights which, by construction, can be understood as a sensible approximation to the preferences of the DMU's. Moreover, we provide an approximate measure of efficiency that depends only on the information related to each DMU, being independent of the rest of the units in the sample. The main drawback of the modified DEA model for DEA users is the calculation of the payoff matrix which usually requires full information about the decision problem that is faced by the DMU's. In a further research, we are working on a way to avoid this difficulty.

References

Allen R, Athanassopoulos A, Dyson R G, Thanassoulis E. Weight restrictions and value judgements in Data Envelopment Analysis: evolution, development and future directions. Annals of Operational Research 1997; 73; 13-34.

Amador F, Sumpsi J M, Romero C. A non-interactive methodology to assess farmers' utility functions: an application to large farms in Andalusia, Spain. European Review of Agricultural Economics 1998; 25; 92-109.

Andersen P, Petersen NC. A procedure for ranking efficient units in DEA. Management Science 1993; 39; 1261-1264.

Arriaza M, Gómez-Limón J A, Upton M. Local water markets for irrigation in Southern Spain: A Multicriteria approach. Australian Journal of Agricultural and Resource Economics 2002; 46; 21-43.

Ballestero E, Romero C. Multiple criteria decision making and its applications to economic problems. Kluwer Academic Publishers: Boston; 1998.

Banker R D, Charnes A, Cooper W W. Some models for estimating technical and scale inefficiencies in DEA. Management Science 1984; 30; 1078-1092.

Bazzani G M. An integrated decision support system for irrigation and water policy design: DSIRR. Environmental Modelling Software 2005; 20; 153-163.

Belton V, Vickers S. Demystifying DEA. A Visual Interactive Approach Based on Multiple Criteria Analysis. Journal of Operational Research Society 1993; 44; 883-896.

Berbel J, Rodríguez A. An MCDM approach to production analysis: An application to irrigated farms in Southern Spain. European Journal of Operational Research 1998; 107; 108-118.

Bouyssou D. Using DEA as a tool for MCDM: some remarks. Journal of Operational Research Society 1999; 50; 974-978.

Cary J W, Holmes W E. Relationships among farmers' goals and farm adjustment strategies: some empirics of a multidimensional approach. Australian Journal of Agricultural Economics 1982; 26; 114-130.

Charnes A, Cooper W W, Golany B, Seiford L, Stutz J. Foundations of DEA for Pareto-

Koopmans efficient empirical production functions. Journal of Econometrics 1985; 30; 91-107.

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

Cooper W W, Seiford L M, Tone K. Data Envelopment Analysis. A comprehensive text with models, applications, references and DEA-Solver Software. Kluwer Academic Publishers: Boston; 2000.

Doyle J, Green R. Data Envelopment Analysis and multiple criteria decision making. Omega 1993; 21; 713-715.

Edwards W. Use of multiattribute utility measurement for social decision making. In: Bell DE, Keeney RL and Raiffa H (Eds) Decisions. John Wiley & Sons: Chichester; 1977.

Farmer P C. Testing the robustness of multiattribute utility theory in an applied setting. Decision Sciences 1987; 18; 178-193.

Francisco S R, Ali M. Resource allocation tradeoffs in Manila's peri-urban vegetable production systems: An application of multiple objective programming. Agricultural Systems 2006; 87; 147–168.

Gasson, R. Goals and values of farmers. Journal of Agricultural Economics 1973; 24; 521-537.

Gómez-Limón J A, Arriaza M, Riesgo L. A MCDM analysis of agricultural risk aversion. European Journal of Operational Research 2003; 151; 569-585.

Gómez-Limón J A, Martínez Y. Multi-criteria modelling of irrigation water market at basin level: A Spanish case study. European Journal of Operational Research 2006; 173; 313-336.

Gómez-Limón J A, Berbel J. Multicriteria analysis of derived water demand functions: a Spanish case study. Agricultural Systems 2000; 63(1); 49-72.

Gómez-Limón J A and Riesgo L. Water pricing: Analysis of differential impacts on heterogeneous farmers. Water Resources Research 2004; 40, Art. No. W07S05.

Hardaker J B, Huirne R B M, Anderson J R. Coping with risk in agriculture. CAB International: Oxon, UK; 1997.

Herath H M G. An empirical evaluation of multiattribute utility theory in peasant agriculture. Oxford Agrarian Studies 1981; 10; 240-254.

Huirne R B M, Hardaker J B. A multi-attribute model to optimize sow replacement decisions. European Review of Agricultural Economics 1998; 25; 488-505.

Joro T, Korhonen P, Wallenius J. Structural Comparison of Data Envelopment Analysis and Multiple Objective Linear Programming. Management Science 1998; 44;

962-970.

Just R E. An investigation of the importance of risk in farmers' decisions. American Journal of Agricultural Economics 1974; 56; 14-25.

Keeney R L, Raiffa H. Decisions with Multiple Objectives: Preferences and Value Trade Offs. John Wiley & Sons: New York;1976.

Markowitz H. Portfolio selection. Journal of Finance 1952; 7; 77-91.

Mas-Colell A, Whinston W D, Green J R. Microeconomic Theory. Oxford University Press: New York; 1995.

Qureshi M E, Harrison S R, Wegener M K. Validation of multicriteria analysis models. Agricultural Systems 1999; 62; 105-116.

Rosen D, Schaffnit C, Paradi C J. Marginal Rates and two dimensional level curves in DEA. Journal of Productivity Analysis 1998; 9; 205-238.

Scheel H. Undesirable Outputs in Efficiency Valuations. European Journal of Operational Research 2001; 132; 400-410.

Smith B, Capstick D F. Establishing priorities among multiple management goals. Southern Journal of Agricultural Economics 1976; 2; 37-43.

Solano C, León H, Pérez E, Herrero M. Characterising objective profiles of Costa Rican dairy farmers. Agricultural Systems 2001; 67; 153-179.

Steward T J. Data Envelopment Analysis and multiple criteria decision making: a response. Omega 1994; 22; 205-206.

Steward T J. Relationships between Data Envelopment Analysis and Multicriteria Decision Analysis. Journal of Operational Research Society 1996; 47; 654-665.

Sumpsi J M, Amador F, Romero C. On Farmers' Objectives: A Multi-Criteria Approach. European Journal of Operational Research 1997; 96; 64-71.

Tone K. A Comparative Study on AHP and DEA. International Journal of Policy Information 1989; 13; 57-63.

Tone K. A Consensus Making Method for Group Decisions. Proposals at the Committee Meeting, National Land Agency: Japan; 1999.

Tone K. A slacks-based measure of efficiency in data envelopment analysis. European Journal of Operational Research 2001; 130; 498-509.

Willock J, Deary I J, Edwards-Jones G, Gibson G J, McGregor M J, Sutherland A, Dent J B, Morgan O, Grieve R. The role of attitudes and objectives in farmer decision making: business and environmentally-oriented behaviour in Scotland. Journal of Agricultural Economics 1999; 50; 286-303.

Young D L. Risk preferences of agricultural producers: their use in extension and research. American Journal of Agricultural Economics 1979; 61; 1063-1070.

FIGURES

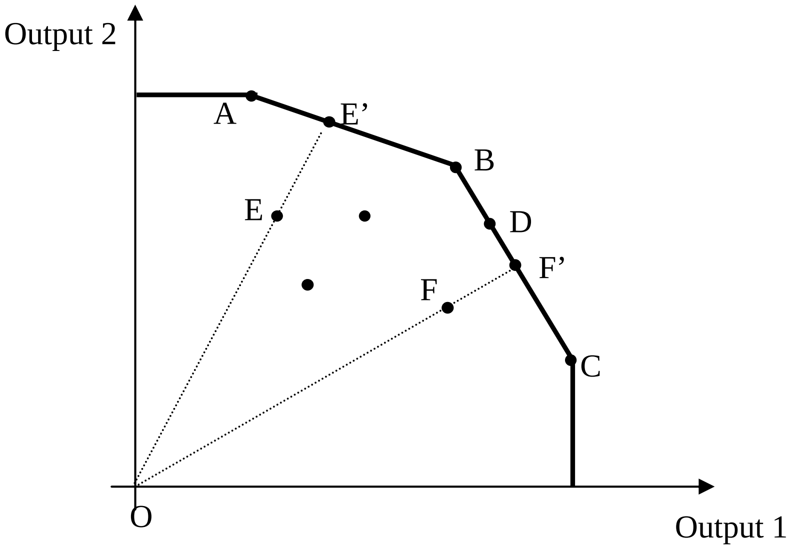

Figure 1. Output oriented BCC model and projection on the frontier

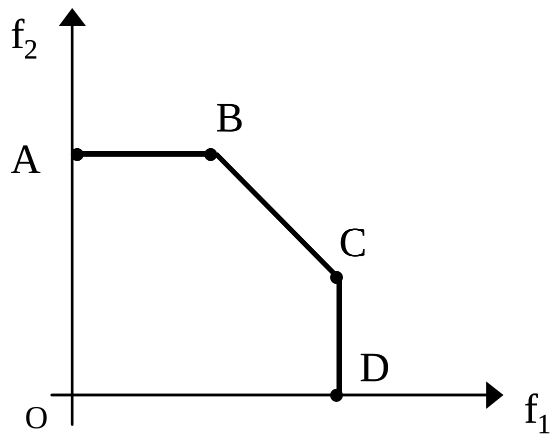

Figure 2. Example of feasible set

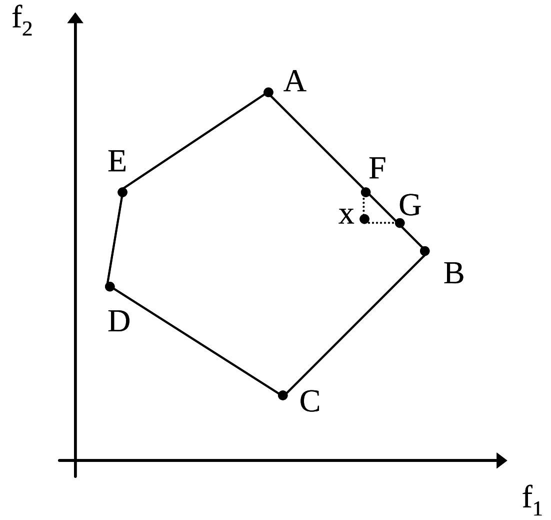

Figure 3. Geometric interpretation of problem [*4*]

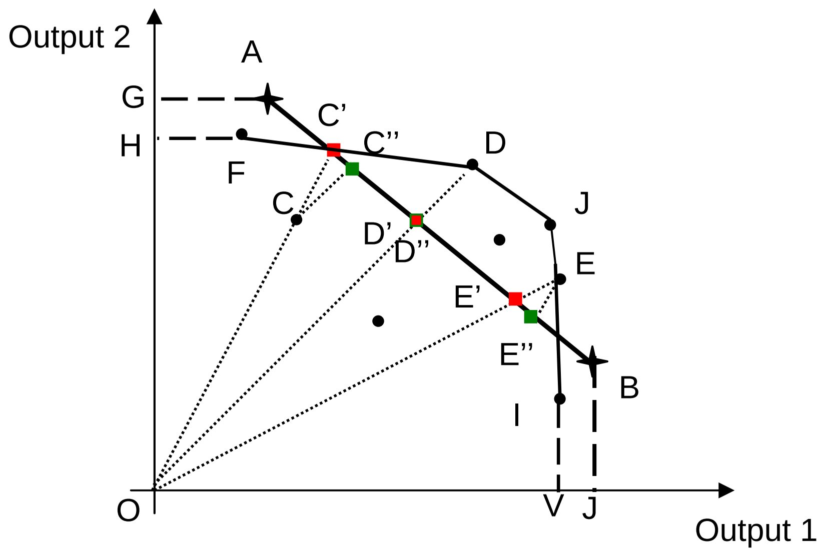

Figure 4. Comparing Sumpsi et al., DEA and modified DEA models

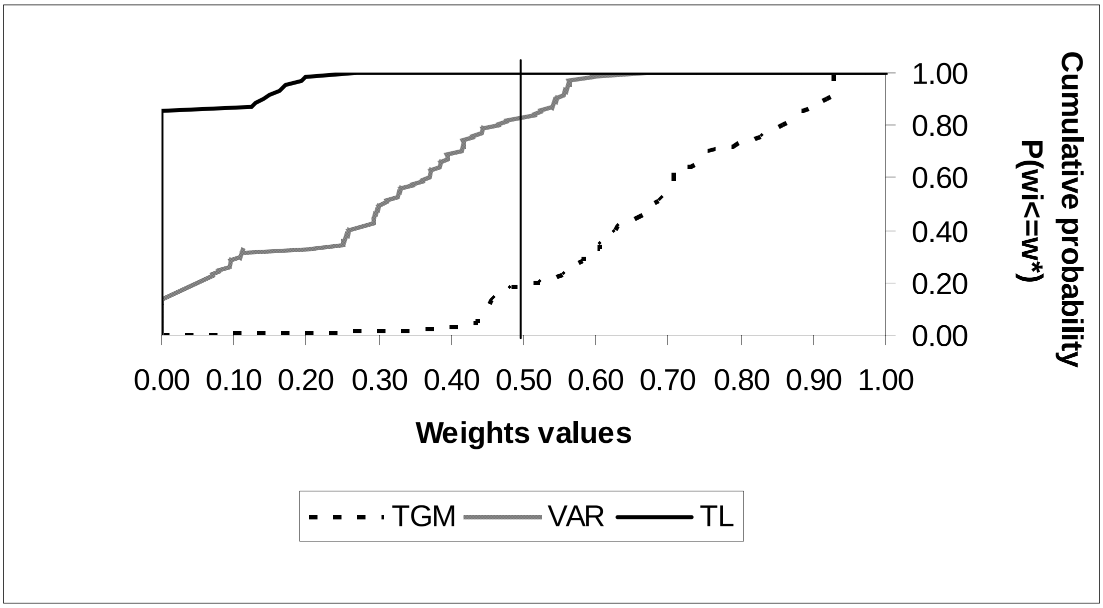

Figure 5. Probability distributions of the weights (w_i) Sumpsi et al. methodology

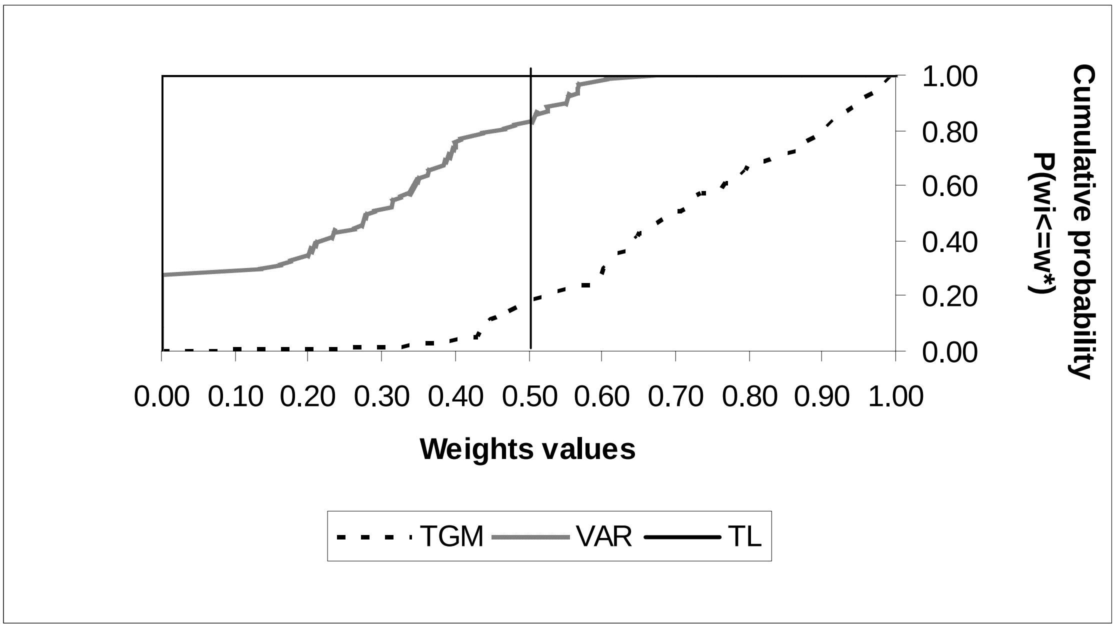

Figure 6. Probability distributions of the weights (λ_i) using modified DEA model

TABLES

Table 1. Basic features of methods

	Required information			*Provided information*		*Projection criterion*
	Observed point	*Payoff matrix*	*Sample*	*Preference weights*	*Efficiency*	
Sumpsi	**X**	**X**		**X**		Min distance
DEA	**X**		**X**		**X**	Radial
M. DEA	**X**	**X**		**X**	**X**	Radial

Table 2. Statistical data about the Sumpsi et al. and modified DEA model weights

	Weights	*Mean*	*Variance*	*Maximum*	*Minimum*	*Median*	*Mode*
Sumpsi et al. approach	TGM	0.681	0.027	0.928	0.328	0.689	0.928
	VAR	0.294	0.037	0.672	0	0.311	0
	TL	0.025	0.004	0.268	0	0	0
Modified DEA approach	TGM	0.729	0.042	1	0.321	0.709	1
	VAR	0.271	0.042	0.679	0	0.291	0
	TL	0	0	0	0	0	0
Mean of the absolute deviation between the weights obtained by both methodologies						TGM	0.051
						VAR	0.034
						TL	0.025

Table 3. Validation using weights estimated by modified DEA method

OBJECTIVES	*Average observed Values*	*Average predicted Values*	*Average deviation*	*Deviation (%)*
TGM (€/ha·year)	1,170.90	1,068.94	*169.21*	*12.60*
VAR (€^{2}/ha·year)	36,302.01	34,511.56	*8,705.02*	*27.85*
TL (hours/ha·year)	35.77	31,77	*4.99*	*15.73*
Decision Variables (ha)	*Average observed crop mix*	*Average predicted crop mix*	*Deviation (ha)*	
Wheat	6.30	16.25	*13.22*	
Maize	82.59	69.51	*15.54*	
Beans	7.08	8.55	*7.47*	
Set-aside	4.18	5.70	*6.15*	

Table 4. Comparing standard DEA and modified DEA model to measure efficiency

	$\bar{\xi}_i$ *(truncated) perturbation*	*Efficiency measure Standard DEA*	*Efficiency measure Modified DEA*
Mean	0.913	0.907	0.950
Standard Error	0.085	0.083	0.075
Correlation with $\bar{\xi}_i$	1.000	0.974	0.827
Correlation with (standard) DEA	0.974	1.000	0.832