

# Value-Based DEA models: application-driven developments

Pedro Nuno de Almeida  
EXIMIUS Consulting  
Rua São Teotónio, Lote 10 - R/Ch Esq., 3000-377 Coimbra, Portugal,  
[pedroalmeida@eximius-consulting.com](mailto:pedroalmeida@eximius-consulting.com)

Luís C. Dias  
INESC Coimbra and Faculty of Economics, University of Coimbra  
Rua Antero Quental 199, 3000-033 Coimbra, Portugal, [lmcdias@fe.uc.pt](mailto:lmcdias@fe.uc.pt)

Data Envelopment Analysis (DEA) is an approach based on linear programming to assess the relative efficiency of peer Decision Making Units (DMUs). Typically, each DMU is free to choose the weights of the factors used in its evaluation. However, the evaluator's preferences may not warrant so much freedom. Several approaches have been proposed to allow the incorporation of managerial preferences in DEA, but few address the Additive DEA model specifically. This paper presents additive DEA models that use Multi-Criteria Decision Analysis concepts to incorporate managerial preferences, and presents the corresponding preference elicitation protocols. The models developed allow the incorporation of preferences at different levels: on valuing performance improvements, on introducing weight restrictions, and on finding adequate targets. These were application-driven developments, resulting from discussing modelling options and preliminary results with the top-level management of a retail chain in the context of an assessment of stores' performance, also described in this paper.

**Key words:** Data envelopment analysis, Multi-criteria decision analysis, additive model, weight restrictions, retail chains.

---

## Introduction

Data Envelopment Analysis (DEA) was devised to consider multiple inputs and outputs in the assessment of the relative performance of Decision Making Units (DMUs) within a network. The pioneering CCR model of Charnes *et al* (1978) assumed constant returns to scale. Later, the BCC model (Banker *et al* 1984) considered variable returns to scale. Each DMU is allowed to choose the weight of each evaluation factor in order to maximize its relative efficiency score. The CCR and BCC models are based on Farrell's radial measure of efficiency, which may indicate a DMU is efficient when in fact it is not truly efficient in Koopman's sense (Färe and Lovell, 1978), thus requiring an analysis of the slacks. The Additive DEA model (Charnes *et al*, 1985) directly identifies Koopmans efficiency by focusing solely on reducing the slacks. The Russell measure (Färe and Lovell, 1978) is another approach to DEA that directly identifies Koopmans efficiency by allowing a non-radial contraction of inputs.

Models such as the Additive DEA model and the Russell measure implicitly assume all factors have equal weight in the production process. The Weighted Additive model (Ali *et al*, 1995) and the Weighted Russell measure (Ruggiero and Bretschneider, 1998) allow this assumption to be dropped. In the latter case, the use of regression techniques has been proposed (Ruggiero and Bretschneider, 1998; Ruggiero, 2000) to derive the weights that best fit

a production process. If price information exists it can be used to measure allocative efficiency (e.g., see Färe et al, 1994, Ch. 7), thus no longer emphasising only technical efficiency.

When focusing on technical efficiency, the client requesting the study may want his or her judgment to influence the allowed trade-offs among inputs and among outputs. Thanassoulis *et al* (2004) set out some of the reasons to constrain the weights allowed for the factors, among them are: to prevent disregarding a factor, to capture prior views on the marginal rates of substitution and/or transformation of the factors of production, to improve discrimination among efficient DMUs, and to ensure that widely differing weights are not assigned to the same factor. Although complete flexibility in specifying trade-offs can be seen as useful to identify inefficiencies (Thanassoulis *et al*, 2004), it might lead to a valuation of inputs and outputs that is inconsistent with the preferences of the client, hindering the acceptability of the results. We should note that the client's subjective value system, subject to the contingencies of the situation (e.g., the need to meet goals), not only may constrain the allowed factor weights but may well intervene earlier in the choice of the set of factors.

Several approaches have been proposed to incorporate the judgment of a client in DEA, especially in CCR and BCC models (see Thanassoulis *et al* (2004) for a review). The most usual approaches are to impose relative weight constraints to define "Assurance Regions", or to incorporate preferences indirectly through changes to the data, whether by transforming the original data vectors (e.g., cone ratio approach), or by including fictitious Unobserved DMUs. Approaches not involving direct inclusion of weight constraints or unobserved DMUs are, for instance, those proposed by Halme *et al* (1999), Podinovski (2004), and Cooper *et al* (2000).

The field of Multi-Criteria Decision Analysis (MCDA) has developed many concepts and protocols to elicit and use the preferences of a client. Bridges between DEA and MCDA have been built by various authors, e.g. Joro *et al* (1998) and Halme *et al* (1999), relating DEA to multi-objective programming, by Bouyssou (1999) and Stewart (1996), relating DEA to MCDA ranking problems, and by Athanassopoulos and Podinovski (1997), relating linear programming formulations used in DEA to those used in MCDA with partial information on weights.

In this paper we build on the bridge between DEA and MCDA built by Gouveia *et al* (2008), which proposed incorporating preferences in the Additive DEA model using concepts from multi-attribute utility/value theory with partial information (e.g. Dias and Clímaco, 2000). Gouveia *et al* suggested converting DEA inputs and outputs into value functions (possibly nonlinear), which are aggregated using the additive MCDA model. The weights are variables to be set by each DMU in order to become the best DMU (if possible) or else to minimize the difference in value to the best DMU. Gouveia *et al*'s paper presented these seminal ideas and illustrated them with an example. As the example used linear value functions there was no need to address the important practical concern of how to elicit value functions in this context. Furthermore, although Gouveia *et al* anticipated that restrictions on the weights could be included in their model, they did not formalize the idea.

We felt the need to develop their methodology in the context of a real-world application, and this raised three important questions. First, how can we elicit a new type of value function that differs in meaning from what is usually

found in MCDA? Second, how can the idea of weight constraints be formalized and how can we elicit such weights? Third, how can we deal with the fact that many alternative optima exist when looking for efficiency targets, many of which do not make sense for the client?

In this paper we present methodological contributions (elicitation protocols, a new formulation with constraints on the weights, and a new formulation for adding constraints to the efficient performance targets) that answer these questions and present the application that motivated them. The application consisted of a small-scale assessment of 19 stores from a retail chain in Portugal, in interaction with the chain's top-level management. An initial model was successively reformulated to reflect management feedback on the results that were being presented. The client organization had previously used output/input ratios to measure productivity. This type of performance assessment has been criticized for yielding only partial productivity measures (not amenable to meaningful aggregation) and for being inadequate to deal with issues such as the age and location of the stores, or the presence of economies of scale (Lusch and Moon, 1984; Kamakura *et al*, 1996). DEA is able to address these concerns, but if we consider specifically the evaluation of units within a chain of retail stores DEA applications are still scarce (as examples see Thomas *et al* (1998), Keh and Chu (2003), Barros and Alves (2003), and Camanho *et al* (2009)).

Despite the application-driven character of the methodological developments, our presentation separates methodology from application. After briefly revisiting Gouveia *et al*'s proposals we present the developed models, and then we present the application that motivated these developments, which at the same time illustrates the proposed models. A concluding section includes some final remarks and suggestions for future work.

### **Incorporation of preferences in Additive DEA: the idea of Value-based DEA**

Charnes *et al* (1985) proposed the Additive DEA model as an alternative to the BCC model, which also considers variable returns to scale. Instead of providing a radial measure of efficiency to be complemented with information about slacks, the Additive model works only with slacks, not requiring a choice between inputs-orientation and outputs-orientation. Ali *et al* (1995) subsequently proposed an Additive model with weights.

Let  $n$  denote the number of DMUs, let  $m$  denote the number of inputs, let  $s$  denote the number of outputs (hence  $q=m+s$  denotes the number of factors), and let subscript  $o$  denote the index of the DMU under evaluation. The linear program to be solved for each  $DMU_o$  according to their model is:

$$\begin{aligned}
 \max_{\lambda, s} z &= \sum_{k=1}^q w_k s_k \\
 s.t.: \quad & \sum_{j=1}^n x_{ij} \lambda_j + s_i = x_{io} \quad (i = 1, \dots, m) \\
 & \sum_{j=1}^n y_{rj} \lambda_j - s_{m+r} = y_{ro} \quad (r = 1, \dots, s) \\
 & \sum_{j=1}^n \lambda_j = 1 \\
 & \lambda_j, s_k \geq 0 \quad (j = 1, \dots, n; k = 1, \dots, q)
 \end{aligned} \tag{1}$$

In (1),  $x_{ij}$  denotes the amount of input  $i$  used by DMU  $j$ ,  $y_{rj}$  denotes the amount of output  $r$  produced by DMU  $j$  and  $(w_1, \dots, w_q)$  is a vector of weights provided beforehand that is fixed for all DMUs. Variables  $\lambda_1, \dots, \lambda_n$  define a convex

combination of the  $n$  DMUs such that in the optimal solution only efficient DMUs will have  $\lambda_j > 0$ . The convex combination corresponds to a projection on the efficient frontier that is better than DMU<sub>o</sub> by a difference  $s_k$  (slack) in each factor  $k$ . Slacks  $s_1, \dots, s_m$  represent reductions in inputs, whereas slacks  $s_{m+1}, \dots, s_q$  represent increases in outputs. This projection is the most distant point from DMU<sub>o</sub> according to the weighted  $L_1$  metric. For instance, in Figure 1 (a case with only two outputs) the inefficient DMU F is enveloped by the efficient frontier ABCDE. If  $w=(1,0)$ , then F is projected into  $F^{(1,0)}$ ; if  $w=(0,1)$  then F is projected into  $F^{(0,1)}$ .

Gouveia *et al* (2008) proposed a value-based methodology to incorporate preferences in Additive DEA, which we will designate as V-DEA, or Value-based DEA. After defining the set of DMUs and the set of inputs and outputs to be used in the evaluation, V-DEA stipulates that outputs and inputs should be converted into value functions to be maximized. Given  $q$  factors (inputs and outputs),  $q$  value functions must be defined such that the worst level has value 0 and the best level has value 1. Hence, after being converted into value all factors are treated as outputs. For each DMU, the value obtained in the multiple factors (which is independent of the original units of measurement) can then be aggregated according to the well-known additive model of MCDA. The aggregation consists of a weighted sum of the values  $v_{1j}, \dots, v_{qj}$  obtained in the multiple factors where the weights  $w_1, \dots, w_q$  represent scaling constants reflecting value trade-offs for the client:

$$v(DMU_j) = \sum_{k=1}^q w_k v_{kj} \quad (2)$$

The V-DEA approach first finds the weights to be used in (2) that most benefit the DMU being analyzed (therefore evaluating it in the best possible light), yielding an efficiency score that can be meaningfully communicated to the client; in a second phase it finds an efficient target in case the DMU is inefficient. The details of this process are as follows:

**Preparatory Phase:** Construction of the  $q$  value functions.

**Phase 1:** For each DMU<sub>o</sub> ( $o=1, \dots, n$ ) a linear program finds the vector of weights  $(w_1, \dots, w_q)$  that minimizes the loss of value to the best DMU when the same vector is used to evaluate all DMUs:

$$\begin{aligned} & \min_{d, w} d \\ & s.t. \quad \sum_{k=1}^q w_k v_{kj} - \sum_{k=1}^q w_k v_{ko} \leq d \quad (j=1, \dots, n) \\ & \quad \sum_{k=1}^q w_k = 1 \\ & \quad w_k \geq 0 \quad (k=1, \dots, q) \end{aligned} \quad (3)$$

By convention, the weights' sum is equal to 1. The score  $d^*$  represents the worst case loss of value or the maximum "regret", a well-known criterion used in decisions with partial information (e.g. Dias and Clímaco, 2000). If  $d^*=0$  in the optimal solution of (3), then DMU<sub>o</sub> is considered to be efficient (although it may be not strictly efficient if some weights are null); otherwise, it is surely inefficient. For the example in Figure 1, the optimal weights for DMU F would be  $w^*=(4/11, 7/11)$ , with  $d^*=0.4$ , for which the DMUs B and C would be the best ones *ex aequo*.

**Phase 2:** For each inefficient DMU<sub>o</sub>, a linear program finds a point in the efficient frontier (the projected efficiency target), considering the optimal vector of weights found for the respective DMU in Phase 1.

$$\begin{aligned}
 \max_{\lambda, s} z &= \sum_{k=1}^q w_k s_k \\
 \text{s.t.} : \quad & \sum_{j=1}^n v_{kj} \lambda_j - s_k = v_{ko} \quad (k = 1, \dots, q) \\
 & \sum_{j=1}^n \lambda_j = 1 \\
 & \lambda_j, s_k \geq 0 \quad (j = 1, \dots, n; k = 1, \dots, q)
 \end{aligned} \tag{4}$$

This linear program corresponds to an oriented Additive DEA model (Ali *et al* 1995) with outputs only. The variables  $\lambda_1, \dots, \lambda_n$  define a convex combination of the  $n$  DMUs. The set of efficient DMUs (possibly only one) defining the convex combination with  $\lambda_j > 0$  are called the “peers” of the DMU<sub>o</sub>. The convex combination corresponds to a point on the efficient frontier that is better than DMU<sub>o</sub> by a difference of value of  $s_k$  (slack) in each factor  $k$ . For the example in Figure 1, the peers of DMU F would be B or C, or both (the linear program allows alternative optimal solutions). The optimal values of (3) and (4) coincide:  $d^* = z^*$ .

V-DEA addresses three shortcomings of Additive DEA: the targets’ dependence on the units of measurement, the pessimistic nature of the projections, and the lack of an intuitive interpretation for the efficiency score. But it should be noted that MCDA has its own limitations: in particular, it requires the assumption of independence among the involved attributes (for details see, e.g., Keeney and Raiffa, 1976), and the questioning procedures underlying the elicitation of value functions and weights require judgments that may be difficult to elicit from decision makers and are subject to biases (e.g., see Pöyhönen and Hämäläinen, 2001).

### **New developments: protocols and extensions for V-DEA**

In the context of a real-world application, we came across three questions that Gouveia *et al* (2008) had not addressed: how the value functions should be interpreted and elicited, how weight constraints could be formalized and elicited, and how the fact that many alternative optima exist for linear program (4) should be dealt with. This section addresses these questions.

#### **Elicitation of value functions**

V-DEA enables the nonlinear conversions of inputs and outputs to a value scale be taken into account, which deals with concerns such as those recently addressed by Cook and Zhu (2009). In V-DEA the purpose of this conversion is to reflect the subjective value assessments of a decision maker (DM). The usual devices of decision analysis can be used (e.g. von Winterfeldt and Edwards 1986; Goodwin and Wright 1998), but the questions must be framed for this particular context. We intend to elicit the difference in the DMU’s merit that corresponds to decreases in inputs or increases in outputs, rather than the value of having these inputs available or outputs produced.

Consider for instance a “Units produced” output. It could happen that after a certain level that corresponds to its natural market these units would need to be sold at a discount (Cook and Zhu 2009). Hence, a DM might consider that if a DMU operating much below that level would see its merit increase  $\Delta v$  by producing  $\Delta y$  more units, another

unit operating above that level would need to increase its production by an amount greater than  $\Delta y$  to obtain the same increase in merit  $\Delta v$ , leading to a concave function typical of MCDA. If a similar output was expressed in terms of “Sales amount” (in currency) then the same DM might consider an increase  $\Delta y$  would be equally meritorious, regardless of the starting level, leading to a linear function. If the output of concern was “market share”, a DM might consider that above a certain level it is much harder to improve market share (hence more meritorious) than if the starting level is very low, leading to a convex function. A convex function might also be appropriate for an input like “number of employees” (if another input was the salary costs). A DM might consider a decrease  $\Delta x$  in the input is more meritorious if a DMU is understaffed than if it is overstaffed. Hence, the value of an understaffed DMU might increase  $\Delta z$  by decreasing  $\Delta x$  employees, but an overstaffed DMU would need to decrease more to obtain the same value increase. By defining convex functions the DM is encouraging larger cuts in inputs that might be abnormally high and larger increases in outputs that are abnormally low (as small variations yield small increases in merit).

Let  $v_k(\cdot)$  denote a value function to be constructed. If the  $k$ -th factor is an input and  $x_{ki}$ ,  $x_{kj}$ ,  $x_{km}$ , and  $x_{kn}$  refer to four consumption levels, then:

- $v_k(x_{ki}) > v_k(x_{kj})$  if and only if using  $x_{ki}$  units of input  $k$  to produce some output reveals more merit (of the DMU) than using  $x_{kj}$  units (normally in this case  $x_{ki}$  would be less than  $x_{kj}$ );
- $v_k(x_{ki}) - v_k(x_{kj}) > v_k(x_{km}) - v_k(x_{kn})$  if and only if there is more merit in reducing consumption from  $x_{ki}$  to  $x_{kj}$  units of input  $k$  than in reducing consumption from  $x_{kn}$  to  $x_{km}$  units, all other things being equal.

By analogy, if the  $k$ -th factor is an output and  $y_{ki}$ ,  $y_{kj}$ ,  $y_{km}$ , and  $y_{kn}$  are four production levels, then:

- $v_k(y_{ki}) > v_k(y_{kj})$  if and only if producing  $y_{ki}$  units of output  $k$  using some input reveals more merit (of the DMU) than producing  $y_{kj}$  units (normally in this case  $y_{ki}$  would be greater than  $y_{kj}$ );
- $v_k(y_{ki}) - v_k(y_{kj}) > v_k(y_{km}) - v_k(y_{kn})$  if and only if there is more merit in increasing production from  $y_{kj}$  to  $y_{ki}$  units of output  $k$  than in increasing production from  $y_{kn}$  to  $y_{km}$  units, all other things being equal.

The elicitation protocol can be based on comparing the merit of increasing an output (or decreasing an input) from  $w$  to  $z$  versus increasing the same output (or decreasing the same input) from  $w'$  to  $z'$ , all other things being equal, and asking the DM to adjust one of these four figures such that the increase of merit would be approximately equal. We will make no assumption about this function except continuity: for instance, it might be partly concave and partly convex (our study shows an example), or it might not even be monotonous (highest value on an ideal level).

### **Incorporation and elicitation of weight constraints**

Weights have a precise meaning in V-DEA: they are scaling coefficients of the value functions that permit marginal value to be converted into global value, enabling possible trade-offs of value between different factors. Directly asking for the weights values should be avoided as the magnitude of the scaling coefficients does not represent the intuitive notion of importance. Rather, they are strongly dependent on the performances chosen to represent the levels 0 and

1 on the value scale. MCDA protocols such as the swings technique (von Winterdeldt and Edwards, 1986; Goodwin and Wright, 1998) are therefore preferable.

Let  $W$  denote a  $q$ -dimensional polyhedron defining the weight vectors the DM finds acceptable, defined by the constraints  $\sum_{k=1}^q w_k = 1, w_k \geq 0$  ( $k=1, \dots, q$ ), plus additional constraints. Additional constraints can include rank order constraints (e.g.  $w_1 \geq w_2$ ), trade-off bounds (e.g.  $w_1 / w_2 \geq 1.5$ ), etc., see Salo and Hämäläinen (2001). In our study the DM felt comfortable indicating a rank order of the weights, plus a limit to the ratio between the weights ranked first and last to avoid null weights. The purpose of this elicitation is only to exclude weights considered unacceptable, while still granting the DMUs freedom to choose their weights. After defining  $W$ , the formulation (3) for Phase 1 has to be slightly adapted to constrain the weights to this set:

$$\begin{aligned} \min_{d, w} & d \\ \text{s.t.} & \sum_{k=1}^q w_k v_{kj} - \sum_{k=1}^q w_k v_{ko} \leq d \quad (j=1, \dots, n) \\ & (w_1, \dots, w_q) \in W \end{aligned} \quad (5)$$

Thus, let  $d^*$  denote the optimal loss of value of DMU<sub>o</sub> relative to the best DMU, obtained for DMU<sub>o</sub>'s optimal weights  $(w_1^*, \dots, w_q^*)$ . We also have to change the formulation (4) for Phase 2; slacks must now be allowed to be negative (see example in Appendix), otherwise it might not be possible to keep the optimal loss of value  $d^*$ :

$$\begin{aligned} \max_{\lambda, s} z &= \sum_{k=1}^q w_k s_k \\ \text{s.t.} & \sum_{j=1}^n v_{kj} \lambda_j - s_k = v_{ko} \quad (k=1, \dots, q) \\ & \sum_{j=1}^n \lambda_j = 1 \\ & \lambda_j \geq 0, \quad s_k \text{ free} \quad (j=1, \dots, n; k=1, \dots, q) \end{aligned} \quad (6)$$

### Model with restrictions on weights and targets

Let  $v^*$  denote the value of the best DMU using the optimal weights found using (5), i.e., the value that DMU<sub>o</sub> ought to achieve, which is such that:

$$v^* = \sum_{k=1}^q w_k^* v_{ko} + d^* \quad (7)$$

The formulation (6) typically admits alternative optimal targets, each of which corresponds to a different way of closing the value gap  $d^*$ . These targets correspond to different projections on the efficient frontier. As we learned from our study, the DM may wish to constrain the efficiency targets that are proposed to achieve  $v^*$ , particularly to avoid those targets that imply an increase of inputs or a decrease of outputs. One idea is simply to impose that value cannot decrease in any factor: targets are forced to maintain or improve performance of all factors.

Let  $S_{<} = \{k \in \{1, \dots, q\} \mid s_k^* < 0 \text{ in Phase 2}\}$  denote the negative slacks in the optimum for (6); these slacks will now become null constants. Let  $S_{\geq} = \{k \in \{1, \dots, q\} \mid s_k^* \geq 0 \text{ in Phase 2}\}$  denote the remaining slacks, which will be

considered as non-negative variables. According to this idea, if  $S_<$  is not empty, a formulation that yields a target alternative to that obtained with (6) can be obtained by solving the following linear program:

$$\begin{aligned}
 & \min_{\{s_k : k \in S_{\geq}\}} \sum_{k=1}^q w_k^* s_k \\
 & s.t.: \sum_{k \in S_{\geq}} w_k^* (v_{ko} + s_k) + \sum_{k \in S_{<}} w_k^* v_{ko} = v^* \\
 & \quad v_{ko} + s_k \leq 1 \quad (k = 1, \dots, q) \\
 & \quad s_k \geq 0 \quad \left( k \in S_{\geq} \right)
 \end{aligned} \tag{8}$$

For each factor, this model will suggest a target value of  $v_{ko} + s_k$  ( $k=1, \dots, q$ ), with  $s_k$  equal to zero if  $k \in S_<$ . Negative slacks are thus avoided, but the target will no longer be a convex combination of the observed DMUs. As in other DEA models, weight constraints are associated with a virtual expansion of the admitted set of production possibilities (Thanassoulis *et al*, 2004). The targets will never exceed value 1 on any factor due to the constraint  $v_{ko} + s_k \leq 1$ . This ensures no extrapolations of the value function are made for ranges outside the elicited interval of performances.

A different possibility is to compute “balanced targets”, i.e., targets that would require improvements in more factors, but less substantial in each factor. The idea is to achieve the target value while minimizing the maximum slack (in terms of value). According to this idea, if  $S_<$  is not empty, a formulation that yields a target alternative to those obtained with (6) and (7) can be obtained by solving the following linear program:

$$\begin{aligned}
 & \min_{\{s_k : k \in S_{\geq}\}} m \\
 & s.t.: \sum_{k \in S_{\geq}} w_k^* (v_{ko} + s_k) + \sum_{k \in S_{<}} w_k^* v_{ko} = v^* \\
 & \quad v_{ko} + s_k \leq 1 \quad (k = 1, \dots, q) \\
 & \quad m \geq s_k \geq 0 \quad (k \in S_{\geq})
 \end{aligned} \tag{9}$$

The next section illustrates the use of these models by presenting the application that motivated their development.

## The motivating application

### Context of the intervention

This section describes a small-scale assessment of stores from a Portuguese retail chain in the pharmacy-cosmetics-hygiene products sector, in which there was interaction with the topmost management level. The methodology proposed by Gouveia *et al* (2008) was suggested to assess the performance of stores within the chain. The client found the possibility of using multiple factors in an integrated model and the prospect of identifying best-practices and identifying sources of inefficiency particularly interesting. This would be useful not only for introducing corrective measures, but also for informing decisions about the configuration and location of future stores. Due to the limitations of the existing information system and also to the agreed time span of the intervention, it was agreed that a small-scale study would be undertaken on one of the sales regions. 19 stores would be selected (the DMUs to analyze) that would cover the entire region while ensuring reasonable homogeneity.



From a discussion with the client's management team (MT) three main interrelated objectives emerged: to indicate which stores are inefficient, to identify typical efficiency profiles, and to propose efficiency targets for the inefficient stores. There was an understanding that the stores' autonomy is limited, to the extent that management is shared between the local supervisor and the chain's managers. On the one hand, the MT makes decisions centrally that affect the stores' performance (such as product mix, supplier negotiations, logistics, pricing, and general human resources policies); on the other hand, it is the local supervisor's responsibility to motivate workers, to maintain customer satisfaction, etc. It was agreed that this study would evaluate this combined management between the chain's MT and the store supervisors.

### **Definition of inputs and outputs**

After defining the objectives and scope of this study, there was a discussion with the MT about the management indicators they regularly used to assess the performance of stores, which were mainly economic-financial ratios. The discussion was geared to highlighting which aspects were being disregarded by those indicators. It also revealed that some of the data that might have been interesting to include in the analysis would probably not be available in time. Furthermore, since there were not many stores to be evaluated, some parsimony in the number of variables was advisable (Dyson *et al* 2001). After discussing several possibilities, a consensus on seven factors emerged:

- *Average stock* (STK): Average value of inventory, in euros (an input);
- *Number of employees* (EMP): Number of full-time equivalent employees (an input);
- *Salary costs* (SAC): Total cost with salaries in euros (an input);
- *Rent* (RNT): Cost of space, in euros (an input);
- *Area* (ARE): Total area of the store, in square meters (an input);
- *Global sales* (SAL): Global value of sales, in euros (an output);
- *Family 4 sales / Global sales* (%F4): Proportion of Family 4 products in global sales (an output).

The average stock (also used by, for instance, Thomas *et al*, 1998; Barros and Alves, 2003), is included for its importance regarding the level of service offered to customers. An inventory matching customers' preferences will minimize investments while minimizing stock-out risks (unfortunately, data on stock-outs was not available). Including the number of employees as an input aims at evaluating the operational productivity of labour, and including the salary costs allows anticipated differences of performance to be taken into account: it is expected that a higher paid salesperson will be a more experienced one who is expected to contribute more to generating sales. Thomas *et al* (1998) and Barros and Alves (2003), for instance, also included these two inputs in their analyses. The remaining two inputs concern the store as real estate. The area is a commonly used input (e.g., Thomas *et al* 1998; Camanho and Dyson 1999; Barros and Alves 2003), as is the rent: Portela and Thanassoulis (2007) and Thomas *et al* (1998) consider that this cost indirectly allows inclusion of the attractiveness of the store's location in the analysis. We should note, however, that it is questionable to mix prices with quantities in technical efficiency assessments. In this case, the fact that all DMUs apply the same salary table to reflect employee experience and the fact that the rents are set

by the property owners (mostly shopping centre developers) as a function of sales expectancies reflecting the store's location, allowed us to take these as adequate proxies for the purposes of this study.

The availability of data limited the choice of output variables. Sales are a commonly used output (e.g.: Barros and Alves, 2003; Athanassopoulos, 2004). The other output chosen was the proportion of sales accounted for by a specific family of products. The client organization is used to monitor the relative weight of the "Family 4" (F4) products as an indicator of customer loyalty. F4 products are expensive and require special customer care and counselling from the salespersons. Therefore, the MT considers that a higher proportion of sales from this family of products indicates a closer proximity to the client and a recognition of a good level of service. According to the MT, it was the proportion of sales (not the absolute sales volume) that would better serve as a proxy.

This set of inputs and outputs was aligned with the concerns and objectives of the MT; it was relevant for all the stores, and it required data that were available and reliable. All these factors were regarded as discretionary from the standpoint of evaluating the combined management of the stores, so that the study would be able to identify efficiency targets related to improvements in factors controlled locally as well as in factors controlled centrally. Table 1 (left) indicates the observed performance of each DMU on the original scale of each factor.

### **Value elicitation for inputs and outputs**

A meeting was held with the MT to learn their views on how variations in the use of inputs and the production of outputs reflected managerial merit, following the developments presented in the "Elicitation of value functions" section. As the goal of "measuring" value accurately is often elusive (von Winterfeldt and Edwards 1986), linear value functions were used whenever this method represented a close approximation to the MT's answers, and piece-wise linear approximations were used for the remaining cases. The elicited ranges were chosen to include the observed performance ranges plus a margin for improvement (or deterioration) of performance.

The resulting value functions linked to a monetary scale (average stock, salaries, rent, and sales) were linear, as shown in Figure 2 (left). The value function for the number of employees is depicted in Figure 2 (centre), and the value function for area is similar. The typical number of employees is in the 12-16 range. Reducing the number of employees from 16 to 12 represents an increase of value (merit), all other things being equal, of 0.15; a reduction by the same number of employees from 20 to 16 is not as valued, since this is seen as a modest reduction for an overstaffed store; in contrast, a reduction from 12 to 8 employees (all other things being equal) represents a lot of merit for the store as the operating conditions become much more difficult. Concerning the proportion of sales from F4 shown in Figure 2 (right), the MT considers that 30% is a normal figure; so there is more merit in increasing one percentage point above that threshold than the same increase below that threshold. But above 40% the value increase is of lesser magnitude, as the MT considers that the store may be reaching an unbalanced sales structure. Note that the value function model is general enough to accommodate a situation where after say, 50%, value might even decrease. Table 1 (right) indicates the result of the conversion of the original units into value.

## Initial models

The first results presented to the MT did not consider constraints on the weights so that a discussion could be fostered about what constraints to introduce. Table 2 gives the results of model (3). Most of the stores are efficient, i.e. there is a vector of weights that makes them the best ones ( $d^*=0$ ). But the chosen vectors include at least two null weights, and for DMU 10 all weights are null except one. As expected, the MT found the existence of these null weights to be undesirable as this corresponded to disregarding some of the factors being used in the evaluation. This clearly motivated the need to introduce weight constraints.

Weight constraints were elicited by asking the MT to compare the “swings” from value 0 to 1 depicted in Table 3. The MT was invited to imagine a store having the value level of 0 for all the factors (first row of Table 3) and to choose the factor where a change to the level of value 1 would represent the highest merit. The MT said that that factor would be Sales, which allows the inference that  $w_{SALES}$  is the highest scaling constant. By repeating this question successively for the remaining factors, the following rank order was elicited:

$$w_{SAL} \geq w_{STK} \geq w_{RNT} \geq w_{SAC} \geq w_{\%F4} \geq w_{EMP} \geq w_{ARE}.$$

To avoid null weights, an indifference judgment question was posed to limit the ratio between the weights ranked first ( $w_{SAL}$ ) and last ( $w_{ARE}$ ): “What would be the highest amount  $z$  that would allow a store with an area of 150 m<sup>2</sup> and annual sales of 500,000 EUR to be considered as having more merit than a store with an area of 450 m<sup>2</sup> and annual sales of  $z$ ?”. The idea was to elicit as high a value  $z$  as possible that was yet low enough to ensure:

$$w_{ARE} v(150 \text{ m}^2) + w_{SAL} v(500,000\text{€}) \geq w_{ARE} v(450 \text{ m}^2) + w_{SAL} v(z)$$

The elicited bound was  $z = 1,000,000 \text{ €}$ , as increasing this amount would make the MT start wondering whether the inequality would hold. Substituting  $z$  in the previous expression yields:

$$w_{SAL} \leq 11.1 w_{ARE}$$

The results for (5) and (6) are presented in Table 4, and indicate only DMUs 2, 10, 12, and 18 as efficient ( $d^*=0$ ). There are no longer any null weights and being the best DMU in some of the factors no longer guarantees being considered efficient. Figure 3 indicates the profiles of the four efficient DMUs, which corresponded to the targets (peers) of the inefficient units. The centre of the heptagon represents value 0 in all factors, whereas the heptagon’s vertices represent value 1. DMU 12 has relatively low value (points near the centre) on inputs (meaning a high level of inputs use) but generates a high value on output sales. The remaining efficient stores have generally higher valued inputs (they use less), but sell less than DMU 12.

Stores 2, 10, and 18 are relatively close in terms of salary costs, but differ substantially in the input number of employees. DMU 18 has more employees but has the lowest average salary per employee, which is approximately 35% below the average salary for DMUs 2 and 10. As the salary per employee is a proxy for the experience of the salespersons, this suggests that experience plays an important role in the performance of the stores; for similar outputs, fewer employees may be required if they are more experienced.

An analogous analysis can be made for the inputs area and rent. The areas of stores 12 and 18 are approximately 50% higher than the areas of stores 2 and 10. As the rent per m<sup>2</sup> of store 12 is approximately twice the rent of store 18, this signals that the location of store 18 is less attractive than the location of store 12. The rent per m<sup>2</sup> of store 18 is also lower than that of the smaller efficient stores 2 and 10. This suggests that for similar outputs, less area may be required if the store has an attractive location.

The efficiency of Stores 2, 10, and 12 agreed with the expectations of the MT. This was an important aspect for the acceptability of the results and less obvious conclusions provided by the models. However, the MT was not expecting DMU 18 to be considered efficient. The explanation is that although DMU 18 sells less than DMU 12, which has a similar area, it compensates for this by having a better performance (more merit) in four inputs (STK, EMP, SAC, and RNT). Compared with DMUs 2 and 10, DMU 18 is worse mainly in the factors that are less important in the weights rank order. DMU 18 is able to offset its relatively large area and high number of employees by not reflecting this scale of operation in the salary costs (inexperienced employees) and rent (less attractive location).

Phase 2 suggests an efficiency target (projection) for each inefficient DMU. The latter would be considered efficient if they changed their value in each factor by the amount indicated in Table 4 (right). But not all the proposed changes correspond to improvements, as some of the slacks are negative (remember that slacks are considered as free variables in (6)). The MT considered that targets suggesting increasing the use of some input or decreasing the production of some output were undesirable, which motivated the development of models (8) and (9).

### **Model with restrictions on weights and targets**

Model (8) was the first to be developed, but the MT considered that the targets were not realistic. Indeed, it happened that the solution suggested that all inefficient DMUs could achieve their target global value  $v^*$  by improving in only one factor but by an unrealistic amount (e.g. decreasing average stock by 70% or increasing sales by 58%, keeping all other factors unchanged). This is due in general to the characteristics of linear programming, but also to the fact that some slacks are set at zero, and the presence of the constraint  $v_{ko} + s_k \leq 1$ . This constraint discourages improvements in factors where the DMU already has a relatively high value and also in factors with low weight (in which a significant increase in value to a figure greater than 1 might be necessary to attain  $v^*$ ).

By discussing these results it was possible to understand that the MT preferred what we later named “balanced targets”, for which model (9) was developed. The new targets are presented in Table 5 in terms of required percentage change on the original units (i.e., the performance change required to obtain the value specified by the slack, according to the value functions, some of which were nonlinear). Although some relatively significant changes remain for some of the DMUs (e.g. the 43.5% decrease in average stock for DMU 8), the MT found no need to place further constraints on the targets. After these improvements the model was now considered as requisite for the client’s purposes.

As an aid to the interpretation of the results a graph similar to the one in Figure 4 (which concerns DMU 1) was drawn for each inefficient DMU. It represents (on the value scales) the performance of the DMU, the performance of

the peer identified in Phase 2, and the value for the balanced target. We can see that the peer DMU was worse than DMU 1 in the sales factor. The alternative target obtains the same global value as the peer in a more balanced way and without worsening the performance in any factor.

## Conclusions

Applying V-DEA in a real-world intervention to assess the performance of stores from a retail chain raised previously unaddressed questions on how to elicit preferences and led to new developments on what concerns constraints on the weights and constraints on the targets. Preference modelling has been used at three different levels. First, input and output measurements were converted into value functions, which required interpreting these value functions as devices to compare the change in merit that corresponds to performance differences. Care was also taken to elicit the value functions for intervals wider than the interval of observed performances. At another level, weight constraints were elicited. The swings technique was partially used to obtain a ranking of the weights, and a trade-off question was used to limit the ratio between the weights ranked first and last. Finally, besides providing Peer DMUs as performance targets for the inefficient stores, two sets of alternative targets were computed: one using a new formulation to avoid negative slacks, another one minimizing the maximum slack (while keeping slacks non-negative). The MT regarded the balanced targets obtained by this process as a realistic complement to the indication of the peer DMUs. The required improvements correspond to adjustments in several factors and do not require radical changes to the production characteristics. Working with value scales both facilitates the incorporation of preferences and lets the Peer DMUs be seen as a model to be attained *in terms of overall value*. This targeted overall value, however, may be attained in different ways.

One interesting aspect of the construction of our models was that it occurred almost in real-time, in direct interaction with the MT, to successively tune the model according to managerial preferences. Naturally, as this study had a limited time span and limited resources, the analysis was not as ambitious as it might otherwise have been if it had been able to involve other stores, other persons (e.g., the store supervisors), or include more factors. The current state of the client's information system particularly hindered the latter desideratum, as many data were not available when needed or at the level of detail required.

The MT deemed the methodology used in this paper as very useful to assess the performance of the retail chain's stores, especially the way the incorporation of preferences enhanced the conclusions obtained. After learning the methodology and the possibilities it offered, the latter were fully explored, e.g. to be able to translate performance into nonlinear value functions or to use weight constraints. One aspect that seems to have been crucial for the MT to learn and to accept the methodology was the constructive character of the model's development, with the results being successively presented and discussed throughout the process so that subsequent changes to the model could be decided.

V-DEA can be applied in many other contexts, and it would be very interesting to test it on new applications which bring new challenges. For instance, although it was not needed in this study, the conversion of the original scale into a value function is a natural way of dealing with undesirable outputs, which are often a problematic aspect (see, e.g. Dyson *et al*, 2001). The future application of V-DEA can further extend the methodology by leaving the DMUs some freedom to shape their own value functions, as suggested by Cook and Zhu (2009) for convex (or concave) functions in the CCR model.

## Acknowledgments

This work has been partly supported by the POCI 2010 FCT grant EGE/58371/2004 and FEDER/COMPETE FCT grant MIT/MCA/0066/2009. The authors wish to express their gratitude to the management team of the client organization for their interest in this study, for their degree of involvement, and for all that we have learned from them. The authors also wish to thank the anonymous referees for their constructive remarks and suggestions.

## Appendix

Figure 5 illustrates why the introduction of weight constraints may yield negative slacks. If DMU G is allowed to freely choose weights in Phase 1, then it would select weights  $w=(0.56,0.44)$  (rounded values) that would minimize its difference in value to the best DMUs with the chosen weights, the DMUs D and E. For instance, its projection could be point  $G^{(0.56,0.44)}$ . However, if a constraint  $w_2 \geq w_1$  is imposed, the previously chosen weights are not admitted and the optimal weights become  $w=(0.5,0.5)$ . Now, the best DMU with the chosen weights is DMU C. Suggesting C as a target for G, however, implies worsening the performance under output  $y_1$  ( $s_1 < 0$ ). If slacks are forced to be positive, then the projection has to be made keeping the value of DMU C when  $w=(0.5,0.5)$ , i.e. in the line joining C and point U. This corresponds to considering an expansion of the production possibilities set or to adding an unobserved DMU U.

## References

- Ali A I, Lerne C S and Seiford L (1995). Components of efficiency evaluations in Data Envelopment Analysis, *Eur J Oper Res* **80**: 462-473.
- Athanassopoulos A D and Podinovski V V (1997). Dominance and potential optimality multiple criteria models decision analysis with imprecise information, *J Oper Res Soc* **48**: 142-150.
- Banker R D, Charnes A and Cooper W W (1984). Some models estimating technical and scale inefficiencies in Data Envelopment Analysis, *Manage Sci* **30**: 1078-1092.
- Barros C P and Alves C A (2003). Hypermarket retail store efficiency in Portugal. *Int J Retail & Distribution Manage* **31**: 549-560.
- Bouyssou D (1999). Using DEA as a tool for MCDM: some remarks, *J Oper Res Soc* **50**: 974-978.
- Camanho A S and Dyson R G (1999). Efficiency, size, benchmarks and targets for bank branches: an application of data envelopment analysis, *J Oper Res Soc* **50**: 903-915.
- Camanho A S, Portela M C and Vaz C B (2009). Efficiency analysis accounting for internal and external non-discretionary factors, *Comput Oper Res* **36**: 1591 – 1601.

- Charnes A, Cooper W W and Rhodes E (1978). Measuring the efficiency of decision making units, *Eur J Oper Res* **2**: 429-444.
- Charnes A, Cooper W W, Golany B, Seiford L and Stutz J (1985). Foundations of Data Envelopment Analysis for Pareto-Koopmans efficient empirical production functions, *J Econometrics* **30**: 91-107.
- Cook W D and Zhu J (2009). Piecewise linear output measures in DEA (third revision), *Eur J Oper Res* **197**: 312-319.
- Cooper W W, Park K S and Pastor J T (2000). Marginal rates and elasticities of substitution with additive models in DEA, *J Prod Anal* **13**: 105-123.
- Dias L C and Clímaco J N (2000). Additive aggregation with Variable Interdependent Parameters: the VIP Analysis software, *J Oper Res Soc* **51**: 1070-1082.
- Dyson R G., Allen R, Camanho A S, Podinovski V V, Sarrico C S and Shale E A (2001). Pitfalls and protocols in DEA, *Eur J Oper Res* **132**: 245-259.
- Färe R and Lovell C A K (1978). Measuring the technical efficiency of production, *J Econ Theory* **19**: 150-162.
- Färe R, Grosskopf S and Lovell C A K (1994). *Production Frontiers*, Cambridge University Press: Cambridge.
- Goodwin P and Wright G (1998). *Decision analysis for management judgment*, Wiley: Chichester.
- Gouveia, M C, Dias L C and Antunes C H (2008). Additive DEA based on MCDA with imprecise information, *J Oper Res Soc* **59**: 54-63.
- Halme M, Joro T, Korhonen P, Salo S E and Wallenius J (1999). A value efficiency approach to incorporating preference information in Data Envelopment Analysis, *Manage Sci* **45**: 103-115.
- Joro T, Korhonen P and Wallenius J (1998). Structural comparison of Data Envelopment Analysis and Multiple Objective Linear Programming, *Manage Sci* **44**: 962-970.
- Kamakura W, Lenartowicz T and Ratchford B (1996). Productivity assessment of multiple retail outlets, *J Retailing* **72**: 325-330.
- Keeney R and Raiffa H (1976). *Decisions with multiple objectives: preferences and value trade-off*, Wiley: New York.
- Keh H T and Chu S (2003). Retail productivity and scale economics at the firm level: a DEA approach, *Omega* **31**: 75-82.
- Lusch R and Moon S (1984). An exploratory analysis of the correlates of labour productivity in retailing, *J Retailing* **67**: 397-419.
- Podinovski V V (2004). Production trade-offs and weight restrictions in data envelopment analysis, *J Oper Res Soc* **55**: 1311-1322.
- Portela M C A S and Thanassoulis E (2007). Comparative efficiency analysis of Portuguese bank branches, *Eur J Oper Res* **177**: 1275-1288.
- Pöyhönen M and Hämäläinen R P (2001). On the convergence of multiattribute weighting methods, *Eur J Oper Res* **129**: 569-585.
- Ruggiero J (2000). Measuring technical efficiency, *Eur J Oper Res* **121**: 138-150.
- Ruggiero J and Bretschneider S (1998). The weighted Russell measure of technical efficiency, *Eur J Oper Res* **108**: 438-451.
- Salo A A and Hämäläinen R P (2001). Preference Ratios in Multiattribute Evaluation (PRIME) - elicitation and decision procedures under incomplete information, *IEEE T Syst Man Cy A* **31**: 533-545.
- Stewart T J (1996). Relationships between Data Envelopment Analysis and Multiple Criteria Decision Analysis, *J Oper Res Soc* **47**: 654-665.
- Thanassoulis E, Portela M C and Allen R (2004). Incorporating value judgments in DEA. In: Cooper, W. W., L. M. Seiford, J. Zhu (eds), *Handbook on Data Envelopment Analysis*, Kluwer, Boston, pp. 99-137.
- Thomas R R, Barr R S, Cron W L and Solum Jr J W (1998). A process for evaluating retail stores efficiency: a restricted DEA approach, *Int J Res Mark* **15**: 487-503.
- von Winterfeldt D and Edwards W (1986). *Decision analysis and behavioural research*, Cambridge University Press: New York.

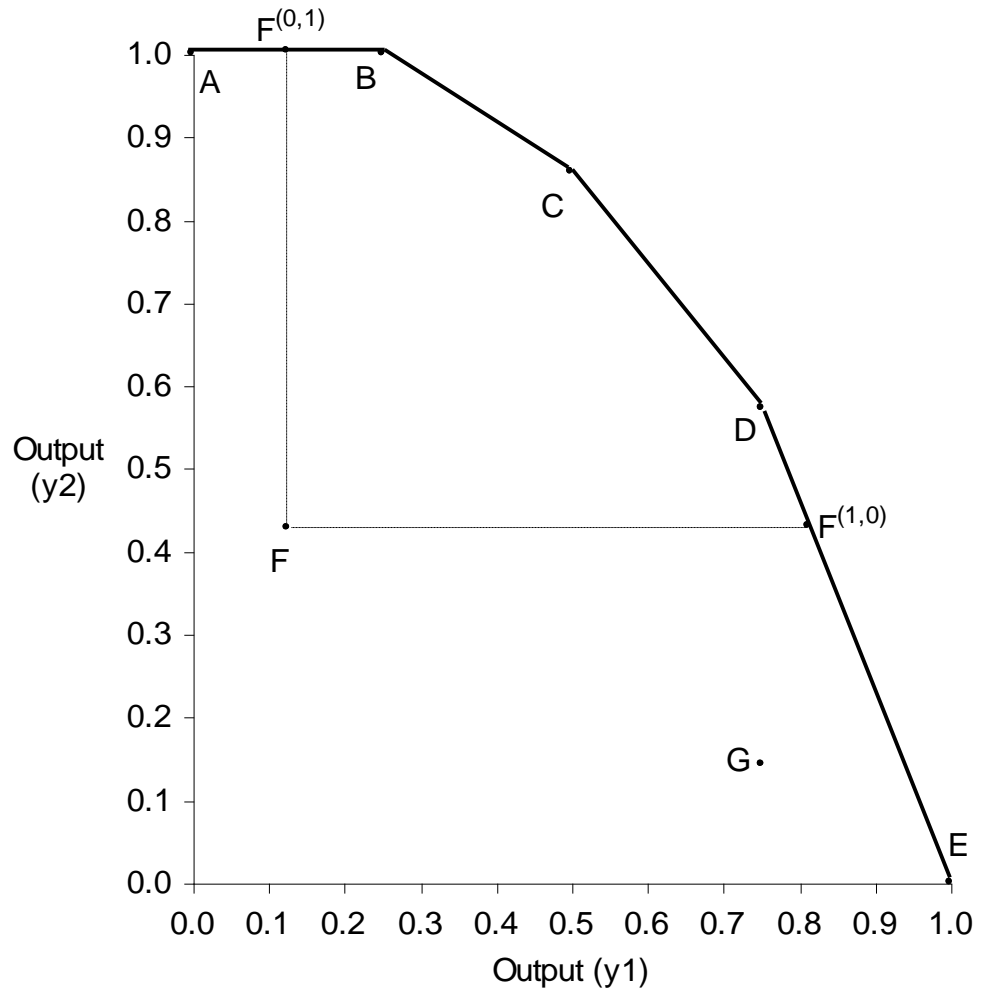


Figure 1 – Projections according to the weighted Additive DEA model.



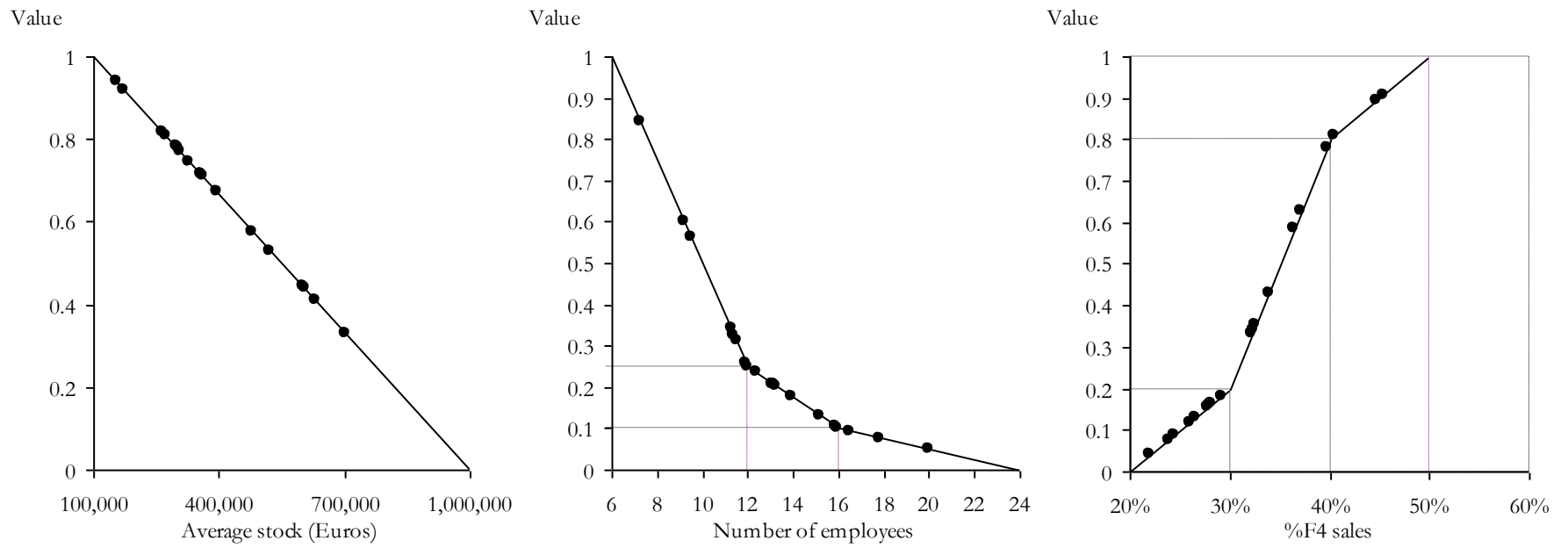


Figure 2 – Three of the value functions elicited (points refer to the observed DMU values).

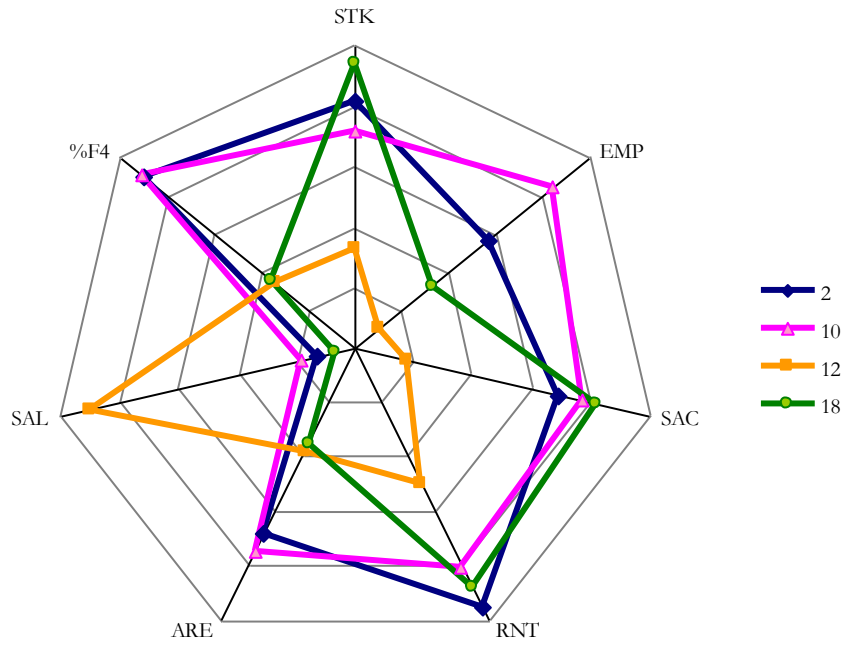


Figure 3 – Profiles for efficient DMUs under weight constraints.

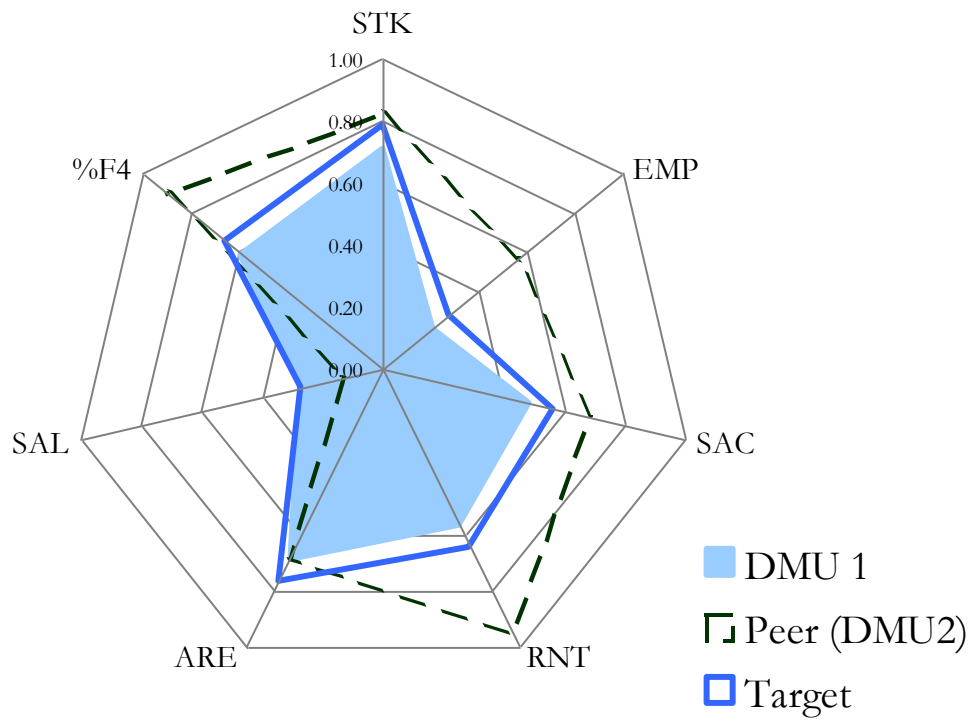


Figure 4 – Observed performance, Peer, and Target for DMU 1 (in units of value).

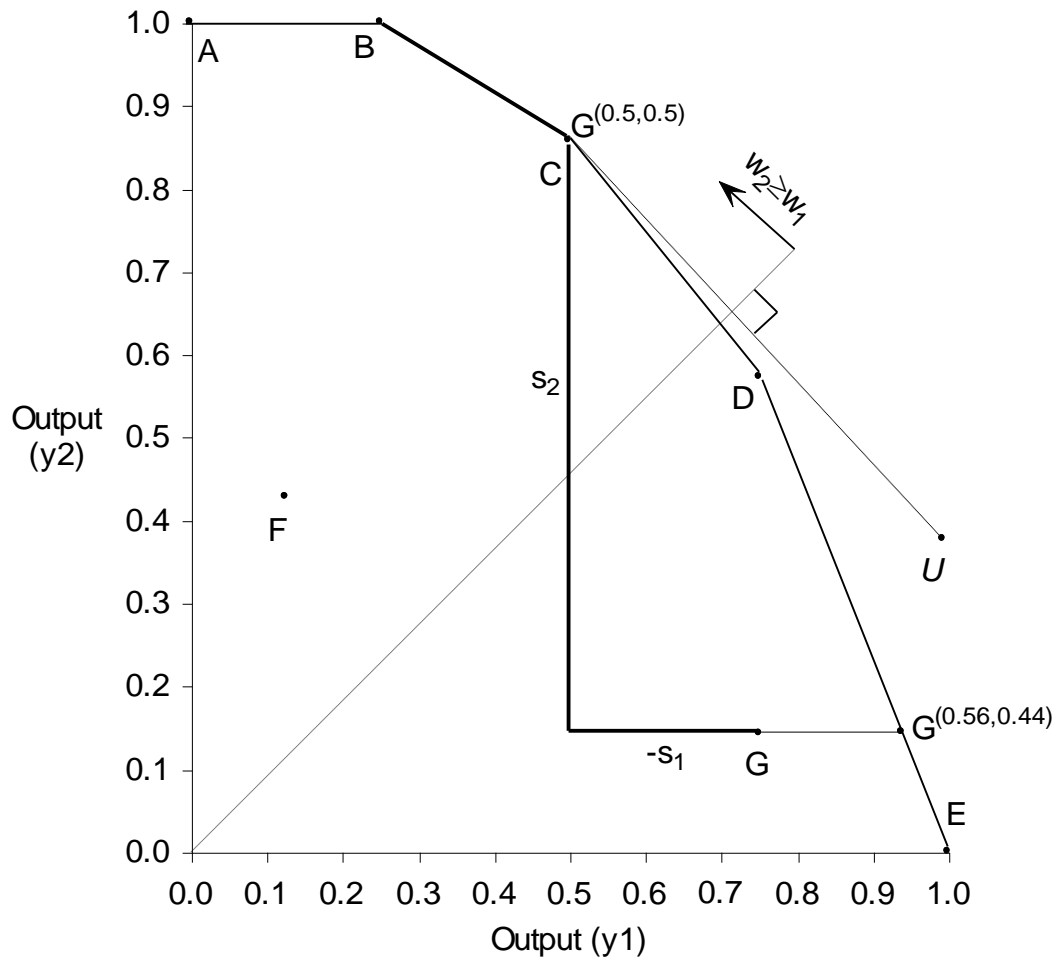


Figure 5 – Expansion of the efficient frontier by weight restrictions or a UDMU.

DMU	Factors (inputs and output) in original scales							Factors in units of value						
	XSTK	XEMP	XSAC	XRNT	XARE	YSAL	Y%F4	VSTK	VEMP	VSAC	VRNT	VARE	VSAL	V%F4
1	360,614	13.2	153,071	275,240	213	1,994,652	36.4%	0.71	0.21	0.48	0.56	0.69	0.27	0.58
2	263,736	9.5	111,409	117,916	213	1,194,289	44.6%	0.82	0.56	0.69	0.96	0.69	0.13	0.89
3	628,938	17.8	218,122	492,305	436	3,841,226	32.2%	0.41	0.08	0.16	0.02	0.03	0.61	0.33
4	479,582	16.5	189,495	134,824	262	2,299,879	23.8%	0.58	0.09	0.30	0.91	0.46	0.33	0.08
5	600,449	15.9	222,567	411,982	331	3,905,880	39.7%	0.44	0.10	0.14	0.22	0.26	0.62	0.78
6	299,876	12.3	159,338	185,368	208	1,554,821	37.2%	0.78	0.24	0.45	0.79	0.71	0.19	0.63
7	171,010	9.2	92,436	124,355	231	625,315	22.0%	0.92	0.60	0.79	0.94	0.60	0.02	0.04
8	354,506	13.9	153,228	231,525	400	1,570,432	24.4%	0.72	0.18	0.48	0.67	0.10	0.19	0.09
9	521,819	13.1	155,918	145,527	222	2,249,522	28.0%	0.53	0.21	0.47	0.89	0.64	0.32	0.16
10	357,204	7.3	96,041	179,931	200	1,505,312	45.4%	0.71	0.84	0.77	0.80	0.75	0.18	0.91
11	307,347	11.3	135,895	171,760	313	1,387,585	28.1%	0.77	0.34	0.57	0.82	0.31	0.16	0.16
12	701,109	15.8	214,814	300,106	290	5,425,809	32.4%	0.33	0.11	0.18	0.50	0.38	0.90	0.34
13	392,894	15.2	170,675	250,726	216	2,269,410	40.5%	0.67	0.13	0.40	0.62	0.67	0.32	0.81
14	604,291	20.0	222,424	387,543	443	3,410,820	27.8%	0.44	0.05	0.14	0.28	0.01	0.53	0.16
15	272,851	12.0	148,268	159,532	197	1,410,839	33.9%	0.81	0.25	0.51	0.85	0.77	0.17	0.43
16	327,304	11.9	181,352	168,006	207	1,263,137	29.1%	0.75	0.26	0.34	0.83	0.72	0.14	0.18
17	356,157	11.5	130,337	181,693	286	1,371,183	26.5%	0.72	0.31	0.60	0.80	0.39	0.16	0.13
18	152,850	11.4	87,223	147,252	301	877,671	32.6%	0.94	0.33	0.81	0.88	0.35	0.07	0.36
19	295,598	13.3	193,606	160,607	199	1,634,121	26.0%	0.78	0.20	0.28	0.85	0.76	0.21	0.12

Table 1 – DMU performances: original scales (left) and value scales (right).

DMU	$d^*$	$w_{STK}^*$	$w_{EMP}^*$	$w_{SAC}^*$	$w_{RNT}^*$	$w_{ARE}^*$	$w_{SAL}^*$	$w_{\%F4}^*$
1	0.007	0.36	0.00	0.04	0.00	0.20	0.36	0.04
2	0.000	0.00	0.17	0.00	0.45	0.00	0.37	0.00
3	0.071	0.53	0.00	0.00	0.00	0.00	0.39	0.08
4	0.000	0.00	0.11	0.00	0.52	0.00	0.37	0.00
5	0.000	0.00	0.23	0.00	0.00	0.00	0.47	0.30
6	0.003	0.47	0.00	0.00	0.00	0.11	0.37	0.04
7	0.000	0.47	0.14	0.00	0.00	0.00	0.39	0.00
8	0.076	0.58	0.00	0.00	0.00	0.00	0.42	0.00
9	0.000	0.00	0.10	0.00	0.50	0.03	0.37	0.00
10	0.000	0.00	0.00	0.00	0.00	0.00	0.00	1.00
11	0.041	0.34	0.00	0.00	0.28	0.00	0.38	0.00
12	0.000	0.00	0.49	0.00	0.00	0.00	0.51	0.00
13	0.000	0.00	0.00	0.00	0.02	0.25	0.40	0.33
14	0.093	0.58	0.00	0.00	0.00	0.00	0.42	0.00
15	0.000	0.50	0.04	0.00	0.00	0.08	0.38	0.01
16	0.035	0.00	0.03	0.00	0.32	0.47	0.18	0.00
17	0.063	0.04	0.00	0.31	0.27	0.00	0.39	0.00
18	0.000	0.49	0.11	0.00	0.00	0.00	0.39	0.01
19	0.000	0.51	0.03	0.00	0.00	0.08	0.38	0.00

Table 2 – Phase 1 results without weight restrictions: optimal distance and weights for each DMU.

Value level	$x_{STK}$	$x_{EMP}$	$x_{SAC}$	$x_{RNT}$	$x_{ARE}$	$y_{SAL}$	$y_{\%F4}$
$v(.)=0$	1,000,000	24	250,000	500,000	450	500,000	20%
$v(.)=1$	100,000	6	50,000	100,000	150	6,000,000	50%

Table 3 – Performances associated to the value levels 0 and 1.

DMU	Phase 1								Phase 2 (slacks in value units)						
	$d^*$	$w^*_{STK}$	$w^*_{EMP}$	$w^*_{SAC}$	$w^*_{RNT}$	$w^*_{ARE}$	$w^*_{SAL}$	$w^*_{\%F4}$	$s^*_{STK}$	$s^*_{EMP}$	$s^*_{SAC}$	$s^*_{RNT}$	$s^*_{ARE}$	$s^*_{SAL}$	$s^*_{\%F4}$
1	0.046	0.38	0.05	0.05	0.05	0.05	0.38	0.05	0.11	0.36	0.21	0.39	0.00	-0.15	0.31
2	<b>0.000</b>	0.21	0.10	0.10	0.21	0.10	0.21	0.10							
3	0.118	0.41	0.04	0.04	0.04	0.04	0.41	0.04	-0.08	0.03	0.02	0.48	0.35	0.29	0.01
4	0.057	0.24	0.03	0.03	0.24	0.03	0.38	0.03	0.24	0.47	0.39	0.04	0.22	-0.20	0.82
5	0.061	0.38	0.03	0.05	0.05	0.03	0.38	0.05	0.37	0.46	0.56	0.74	0.43	-0.49	0.11
6	0.035	0.38	0.03	0.03	0.10	0.03	0.38	0.03	0.04	0.32	0.24	0.17	-0.02	-0.07	0.26
7	0.003	0.23	0.02	0.23	0.23	0.02	0.23	0.02	0.02	-0.27	0.03	-0.06	-0.25	0.05	0.32
8	0.106	0.38	0.03	0.07	0.07	0.03	0.38	0.03	0.10	0.38	0.21	0.28	0.59	-0.07	0.80
9	0.060	0.24	0.03	0.03	0.24	0.03	0.38	0.03	0.29	0.35	0.22	0.07	0.05	-0.19	0.73
10	<b>0.000</b>	0.20	0.08	0.20	0.20	0.02	0.20	0.08							
11	0.067	0.38	0.03	0.07	0.07	0.03	0.38	0.03	0.05	0.22	0.12	0.13	0.37	-0.04	0.73
12	<b>0.000</b>	0.10	0.10	0.10	0.10	0.10	0.41	0.10							
13	0.034	0.38	0.03	0.05	0.05	0.03	0.38	0.05	0.14	0.43	0.30	0.33	0.02	-0.20	0.08
14	0.137	0.38	0.05	0.05	0.05	0.03	0.38	0.05	-0.11	0.06	0.04	0.22	0.37	0.37	0.19
15	0.029	0.38	0.03	0.03	0.10	0.03	0.38	0.03	0.01	0.31	0.18	0.10	-0.08	-0.04	0.46
16	0.080	0.38	0.03	0.03	0.10	0.03	0.38	0.03	0.07	0.30	0.35	0.13	-0.03	-0.01	0.71
17	0.089	0.38	0.03	0.07	0.07	0.03	0.38	0.03	0.10	0.25	0.09	0.16	0.29	-0.03	0.76
18	<b>0.000</b>	0.23	0.02	0.23	0.23	0.02	0.23	0.02							
19	0.044	0.38	0.03	0.03	0.10	0.03	0.38	0.03	0.04	0.36	0.41	0.11	-0.07	-0.08	0.77

Table 4 – Phase 1 and Phase 2 results under weight constraints and free slacks.

DMU	$\Delta X_{STK}$	$\Delta X_{EMP}$	$\Delta X_{SAC}$	$\Delta X_{RNT}$	$\Delta X_{ARE}$	$\Delta y_{SAL}$	$\Delta y_{\%F4}$
1	-18.4%	-10.7%	-9.6%	-10.7%	-6.9%		3.4%
3		-33.8%	-18.2%	-16.1%	-21.8%	28.4%	10.3%
4	-17.8%	-17.3%	-10.0%	-25.8%	-9.1%		19.9%
5	-14.9%	-16.6%	-8.9%	-9.6%	-10.0%		10.9%
6	-17.8%	-5.7%	-7.4%	-12.8%			2.7%
7	-2.4%		-1.0%			4.0%	1.0%
8	-43.5%	-19.5%	-22.4%	-29.6%	-18.4%		26.8%
9	-16.6%	-11.7%	-12.4%	-26.5%	-8.7%		10.5%
11	-32.0%	-7.8%	-16.1%	-25.5%	-11.6%		10.9%
13	-12.7%	-9.7%	-6.5%	-8.8%	-5.1%		6.8%
14		-40.9%	-19.9%	-22.9%	-23.7%	35.8%	18.7%
15	-16.6%	-3.3%	-6.8%	-12.6%			2.5%
16	-37.9%	-9.2%	-15.2%	-32.8%			9.9%
17	-36.3%	-10.0%	-22.0%	-31.6%	-15.1%		17.7%
19	-23.1%	-11.2%	-7.8%	-18.9%			14.6%

Table 5 – Balanced targets: required variation in the original units.