Journal of Productivity Analysis, 16, 241–261, 2001.

Lower Bound Restrictions on Intensities in Data Envelopment Analysis

SYLVAIN BOUHNIK, BOAZ GOLANY AND URY PASSY
Faculty of Industrial Engineering and Management, Technion—Israel Institute of Technology, Haifa 32000, Israel

STEVEN T. HACKMAN
School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

DIMITRA A. VLATSA
Experian Inc., Atlanta, GA, USA

Abstract

We propose an extension to the basic DEA models that guarantees that if an intensity is positive then it must be at least as large as a pre-defined lower bound. This requirement adds an integer programming constraint known within Operations Research as a Fixed-Charge (FC) type of constraint. Accordingly, we term the new model DEA_FC. The proposed model lies between the DEA models that allow units to be scaled arbitrarily low, and the Free Disposal Hull model that allows no scaling. We analyze 18 datasets from the literature to demonstrate that sufficiently low intensities—those for which the scaled Decision-Making Unit (DMU) has inputs and outputs that lie below the *minimum* values observed—are pervasive, and that the new model ensures fairer comparisons without sacrificing the required discriminating power. We explain why the "low-intensity" phenomenon exists. In *sharp* contrast to standard DEA models we demonstrate via examples that an inefficient DMU may play a pivotal role in determining the technology. We also propose a goal programming model that determines how deviations from the lower bounds affect efficiency, which we term the trade-off between the deviation gap and the efficiency gap.

Keywords: data envelopment analysis, goal programming

1. Introduction

In today's competitive business world efficiency measurement has received increased attention from managers in all sectors of the economy. Evaluating a firm's performance relative to the other firms within an industry, and identifying sources of inefficiency are the first steps in any process of improvement. Since its introduction in 1978, Data Envelopment Analysis (DEA) has been the most widely used nonparametric, activity analysis methodology to estimate the relative efficiency of similar firms or Decision-Making Units (DMUs) within an organization (Seiford, 1995). DEA estimates a DMU's efficiency by determining how far it deviates from the *best-practice* frontier consisting of *composite* DMUs, which are optimally-weighted linear combinations of DMUs in the dataset.

Each DMU_j, $j = 1, 2, \ldots, N$, uses X_{ij} amount of resource (input) $i, i = 1, 2, \ldots, m$, to produce or provide Y_{rj} amount of product or service (output) $r, r = 1, 2, \ldots, p$. It is assumed that all inputs and output quantities are nonnegative, and that the input–output vector $(X_j, Y_j) \in R_+^m \times R_+^p$ associated with each DMU_j is not identically zero. The linear programming formulation used for computing the *input efficiency* of DMU_k is summarized below:

DEA input efficiency model

$$\text{Min } \theta_k \tag{1}$$

$$\sum_{j=1}^{N} \lambda_j Y_{rj} \geq Y_{rk}, \quad r = 1, \ldots, p \tag{2}$$

$$\sum_{j=1}^{N} \lambda_j X_{ij} \leq \theta_k \cdot X_{ik}, \quad i = 1, \ldots, m \tag{3}$$

$$\lambda = (\lambda_1, \ldots, \lambda_N) \in \Lambda \tag{4}$$

The variable λ_j is commonly referred to as the *intensity* associated with DMU_j. The solution λ^* and θ_k^* to the linear program has the following interpretation. There exists a (hypothetical) composite DMU formed as a nonnegative linear combination of existing DMUs using the components of λ^*. This composite DMU produces at least as much output as DMU_k and consumes no more than $100\theta^*\%$ of each input of DMU_k. From this perspective DMU_k can theoretically scale all of its inputs by $100(1 - \theta^*)\%$ amount and still achieve its current levels of output. When $\Lambda = R_+^N$, equations (1–4) define the *Constant Returns to Scale* (CRS) model (Charnes et al., 1978). When $\Lambda = \{\lambda \in R_+^N : \sum_{j=1}^{N} \lambda_j = 1\}$, it define the *Variable Returns to Scale* (VRS) model (Banker et al., 1984). Other model formulations, including some that allow nonconvex reference technologies have since been suggested (see, e.g., Petersen, 1990).

Like all extrapolation techniques the accuracy of DEA's efficiency rating and its acceptance by management is critically dependent on how well the constructed efficiency frontier corresponds to the true efficiency frontier. An essential question should always be: *How reasonable is the composite DMU to which the reference DMU is being compared?* The research reported herein was motivated by a recent application of DEA to benchmarking the warehouse and distribution industry (Frazelle and Hackman, 1993; Hackman et al., 2001; Vlatsa, 1995). When using the CRS and VRS models, quite often intensities lower than 2% were observed, which means that a component warehouse of a composite warehouse was scaled to less than 2% of its original size. Almost 80% of the inefficient warehouses had composite warehouses made up of at least one warehouse scaled to less than 10%. In many cases the scaled component warehouse was smaller than any actual warehouse in the dataset. In one extreme example a component warehouse was used at a scale of only 0.04%, which translated into a 92 square foot facility that would use 205 total labor hours (1 person for 5 weeks) and would invest \$1,632 in material handling and storage equipment! The infrastructure and operations of a warehouse are designed for a particular type of activity, which limit the extent to which the "blueprint" may be scaled downwards. Using low intensities in the construction of a composite DMU is certainly not realistic for this industry from either an economic or modeling perspective. Equally as important, it is our experience that managers often question the meaning of convex combinations that involve what they perceive to be irrelevant DMUs.

One way to eliminate this problem is to argue that comparisons are only reasonable if they are made with respect to observed rather than composite units, namely, insist that $\Lambda = \{\lambda \in R_+^N : \lambda_j \in \{0, 1\}\}$. This type of restriction is the *Free Disposable Hull* (FDH) model presented in Tulkens (1993). However, in our experience, the FDH model tends to assess most DMUs as efficient, which may limit its usefulness. Another way to handle the low intensities is to sequentially run DEA, each time forcing $\lambda_j = 0$ for each DMU$_j$ that received a sufficiently low intensity value in the previous run. While this scheme is possible, it is computationally cumbersome, nonoptimal, and can be avoided by replacing the linear model with a mixed-integer model we shall propose here.

The present paper takes a *middle-of-the-road* approach that lies between the two extremes—the original DEA models that allow DMUs to be scaled arbitrarily low and the FDH model that allows *no* scaling. In our proposed approach, composite units can be constructed by taking any linear combination of DMUs provided that the contribution of each component DMU that participates in the construction is not negligible. The proposed model(s) restrict intensities to appropriate (nonconvex) regions that are designed to ensure fair comparisons *without sacrificing the required discriminating power* inherent in the CRS and VRS DEA models. Specifically, in the CRS case if λ_j is positive then it must be at least as large as a predefined *lower bound* ℓ_j. The set Λ is thus constrained to lie within the set $\{\lambda \in R_+^N : \lambda_j \in \{0\} \cup [\ell_j, \infty)\}$. Since this restriction leads to an integer-programming constraint known within the operations research parlance as a *Fixed-Charge* constraint, we shall term this new model as CRS_FC or VRS_FC, depending on whether the constant or variable returns-to-scale model is used. To the extent that upward scaling of smaller units lead to unrealistic results, as argued by Thore (1996), then one should add corresponding upper bound constraints, which will not appreciably affect the computational time with today's commercial solvers. In this paper we have chosen to focus on lower bounds for two reasons: (1) scaling upwards can often be justifiable using the replication argument, and so it is less egregious, and (2) upper bounds are unnecessary for the VRS models.

Attempts to restrict weights in DEA have been recorded since the mid-80s. Most of these efforts concentrated on restricting weights attributed to input-output factors in the primal (the *virtual multiplier*) formulation of DEA (see Dyson and Thanassoulis, 1988; Charnes et al., 1989; Roll et al., 1991). Restricting the intensities in the dual, the *envelopment* formulation of DEA, has attracted attention only in the last few years. In applying DEA to evaluate the performance of computer companies, Thore (1996) observed that some of the largest companies (e.g., IBM) were rated as inefficient and their composite DMUs were constructed by some of the smallest companies in that market represented by huge intensities. Thore introduced constraints to exclude such cases from his analysis. Golany and Thore (1997) also developed models that restricted the envelopment intensities, such as preventing a particular DMU's intensity from being larger than another DMU, or forcing zero weights on units that fall under a certain threshold.

The paper is organized as follows. In Section 2 we demonstrate the pervasive nature of the *low-intensity* phenomenon by analyzing 18 previous DEA applications in which data was included. In Section 3 we formulate the Fixed Charge (DEA_FC) model, and implement the new model with 3 different choices for lower bounds on the same datasets analyzed in Section 2. The results demonstrate that while specific DMUs may see their efficiency scores dramatically altered, the overall efficiency scores are reasonably close to the original DEA models. In Section 4 we describe the mathematical and geometrical properties of CRS_FC. We begin by explaining why the low intensity phenomenon exists. Next, in *sharp* contrast to the convex DEA models we show via simple examples that in *nonconvex* models an inefficient DMU may be part of the composite DMU and that a scaled version of an inefficient DMU may be efficient. We close the section by describing the inherent nonconvexities of the new model and illustrate the resultant technologies through a graphical construction of the level sets for a 2-input, scalar output case. In Section 5 we present a class of goal programming extensions to the proposed model that trade-off what we term the *efficiency gap* with the *deviation gap*. That is, as the constraint on lower bounds is loosened the corresponding efficiency score is lowered, which permits an assessment of the robustness of the DEA_FC model.

2. Analysis of Existing Datasets

To assess the pervasiveness of *low-intensities,* in the absence of previous results, we decided to define a common criterion and then apply it to the analysis of eighteen existing datasets. Let Min_DMU denote a hypothetical DMU whose input-output vectors (X^{Min}, Y^{Min}) are the minimum over all observed inputs and outputs, namely, $X_i^{Min} = \min_j X_{ij}$ and $Y_r^{Min} = \min_j Y_{rj}$. Let λ^* denote an optimal solution obtained for the standard (CRS or VRS) DEA model. Conceivable rules for determining a uniform lower bound on the various intensities are given by:

$$
\begin{aligned}
\ell_j^{in} &= \min_i \left(X_i^{Min} / X_{ij} \right) \\
\ell_j^{out} &= \min_r \left(Y_r^{Min} / Y_{rj} \right) \\
\ell_j &= \min \left\{ \ell_j^{in}, \ell_j^{out} \right\}.
\end{aligned}
\tag{5}
$$

Let λ^* denote an optimal solution obtained for the standard (CRS or VRS) DEA model. A composite DMU (for an inefficient DMU) is said to exhibit a *strong violation* if $\lambda_j^* < \ell_j$; exhibit a *strong input violation* if $\lambda_j^* < \ell_j^{in}$; or exhibit a *weak violation* if either $\lambda_j^* X_{ij} < X_i^{Min}$ for at least one input i or $\lambda_j^* Y_{rj} < Y_r^{Min}$ for at least one output r. Obviously, a strong violation also counts as a weak violation. A composite DMU that has a component with a type of violation is said to exhibit that same violation, and for the present analysis is not considered a valid composite DMU. We begin by examining the dataset describing physician's hospital practices presented in Chilingerian (1995) and reproduced in Table 1. The optimal intensities

Table 1. Results of the efficiency analysis (VRS model) (Data from Chilingerian, 1995: Group 2, Table 9-3).

Data					Results						
X_1	X_2	Y_1	Y_2	DMU	λ_2	λ_3	λ_4	λ_7	λ_{10}	λ_{12}	Efficiency
227	102330	14	20	DMU1	0.000	0.000	0.350	0.190	0.00	0.450	0.97885
132	105320	26	12	DMU2	1.00	—	—	—	—	—	1.00
348	342657	47	18	DMU3	—	1.00	—	—	—	—	1.00
260	106876	7	27	DMU4	—	—	1.00	—	—	—	1.00
380	161349	17	23	DMU5	0.200	0.067	0.510	0.000	0.220	0.000	0.80893
344	237264	14	21	DMU6	0.400	0.000	0.600	0.000	0.000	0.000	0.60698
243	92890	12	19	DMU7	—	—	—	1.00	—	—	1.00
290	118463	12	24	DMU8	0.110	0.000	0.700	0.000	0.098	0.098	0.92343
241	103467	14	15	DMU9	0.000	0.000	0.000	0.630	0.000	0.370	0.91654
558	143122	23	25	DMU10	—	—	—	—	1.00	—	1.00
367	151373	18	21	DMU11	0.360	0.009	0.390	0.000	0.230	0.000	0.77269
184	98071	21	15	DMU12	—	—	—	—	—	1.00	1.00
132	92890	7	12	Min_DMU							

with the efficiency scores (six DMUs are not efficient) are given, and the coordinates of Min_DMU are also shown.

Table 2 shows the corresponding scaled DMUs comprising the component and composite DMUs for the inefficient DMUs. The rows highlighted in bold correspond to component DMUs that exhibited a strong violation. With respect to the composite DMUs: 4 of the 6 cases had at least one component that exhibited a strong violation, and *all* 6 cases had at least one component that exhibited either a strong input or weak violation. With respect to the 19 component DMUs: 11 exhibited a strong violation, 15 exhibited a strong input violation, and *all* 19 exhibited a weak violation!

To assess the extent of the low-intensity problem, for each of the selected 18 datasets, we adopted the Min_DMU criterion. Table 3 presents the empirical findings for the VRS and CRS models, respectively. The low-intensity problem is indeed pervasive. Consider first the VRS model. With respect to composite DMUs corresponding to inefficient DMUs: 68% (202 out of 297) had at least 1 component exhibiting a strong violation, 82% (244 out of 297) had at least one component exhibiting a strong input violation, and 98% (290 out of 297) had at least one component exhibiting a weak violation. With respect to component DMUs: almost 3 in 10 exhibited a strong violation, 4 in 10 exhibited a strong input violation, and 3 in 4 exhibited a weak violation. The severity of the low-intensity problem for the CRS model is much less than the VRS model, but still significant. We conclude from our empirical analysis that very often both the VRS and CRS models construct composite DMUs with small enough components that may be viewed as unreasonable *depending on the particular application*.

We emphasize that for certain applications, most notably in which resources are easily scalable (e.g., labor), adding lower bound restrictions may be unnecessary. For other applications, most notably those that involve capital investments that only make sense above a minimum level, adding lower bound restrictions is a simple way to avoid unreasonable composites.

Table 2. Composite DMUs with the VRS model (Data from Chilingerian, 1995: Group 2, Table 9-3).

Inefficient DMU	Component DMUs in Composite DMU					Violation	No.
		X_1	X_2	Y_1	Y_2		
DMU_1	$\lambda_4 DMU_4$	**91.52**	**37620**	**2.464**	**9.504**	Strong	1
	$+\lambda_7 DMU_7$	**47.142**	**18020**	**2.328**	**3.686**	Strong	2
	$+\lambda_{12} DMU_{12}$	83.536	44524	9.534	6.81	Weak	3
	= Composite DMU	222.198	100164	14.326	20		
DMU_5	$\lambda_2 DMU_2$	**25.942**	**20698**	**5.1098**	**2.3584**	Strong	4
	$+\lambda_3 DMU_3$	**23.387**	**23028**	**3.1586**	**1.2097**	Strong	5
	$+\lambda_4 DMU_4$	133.29	54790	3.5886	13.842	Weak	6
	$+\lambda_{10} DMU_{10}$	**124.78**	**32004**	**5.1431**	**5.5903**	Strong	7
	= Composite DMU	307.399	130520	17	23		
DMU_6	$\lambda_2 DMU_2$	52.8	42128	10.4	4.8	Weak	8
	$+\lambda_4 DMU_4$	156	64126	4.2	16.2	Weak	9
	= Composite DMU	208.8	106254	14.6	21		
DMU_8	$\lambda_2 DMU_2$	**14.368**	**11464**	**2.83**	**1.3062**	Strong	10
	$+\lambda_4 DMU_4$	180.88	74354	4.8699	18.784	Weak	11
	$+\lambda_{10} DMU_{10}$	**54.578**	**13999**	**2.2496**	**2.4452**	Strong	12
	$+\lambda_{12} DMU_{12}$	**17.966**	**9575.8**	**2.0505**	**1.4646**	Strong	13
	= Composite DMU	267.792	109392	12	24		
DMU_9	$\lambda_7 DMU_7$	151.92	58075	7.5024	11.879	Weak	14
	$+\lambda_{12} DMU_{12}$	68.964	36757	7.8709	5.622	Weak	15
	= Composite DMU	220.884	94832	15.3733	17.501		
DMU_{11}	$\lambda_2 DMU_2$	47.961	38267	9.4468	4.3601	Weak	16
	$+\lambda_3 DMU_3$	**3.2917**	**3241.2**	**0.44457**	**0.17026**	Strong	17
	$+\lambda_4 DMU_4$	**102.65**	**42196**	**2.7637**	**10.66**	Strong	18
	$+\lambda_{10} DMU_{10}$	**129.67**	**33260**	**5.3449**	**5.8097**	Strong	19
	= Composite DMU	283.572	116964	18	21		
	Min_DMU	132	92890	7	12		

3. The Fixed Charge Model: Development and Empirical Comparison with Other DEA Models

We begin by formally defining the empirical production possibility set

$$T^{FC} = \{(X, Y) \in R_+^m \times R_+^p : X \text{ can produce } Y\} \tag{6}$$

associated with a fixed charge (FC) technology. We shall find it instructive to first reexamine the CRS case. Recall that for the CRS technology the empirical production possibility set that is generated by a set of observed data $(X_j, Y_j),\ j = 1, 2, \ldots, N$ is algebraically defined as (see Banker et al., 1984, pp. 1081–1082)

$$T^{CRS} = \left\{(X, Y) \in R_+^m \times R_+^p : X \geq \sum_j \lambda_j X_j,\ Y \leq \sum_j \lambda_j Y_j \text{ for some } \lambda \in R_+^N\right\}. \tag{7}$$

Table 3. Efficiency changes from BCC model to BCC model with restrictions: BCC ⟹ weak restrictions, BCC ⟹ strong restrictions, BCC ⟹ FDH model.

BCC Model		Efficiency					
Efficiency	%	E	90–99	80–89	70–79	60–69	<60
		Lower Weak Bounds %					
E	48.267	48.267					
90–99	18.911	5.7988	13.112				
80–89	10.809	0.000	1.8334	8.9752			
70–79	8.460	0.2646	0.000	0.9006	7.2951		
60–69	7.352	0.000	0.1587	0.000	0.9848	6.2082	
<60	6.202	0.1587	0.000	0.1587	0.000	0.1984	5.6861
Sum	100	54.489	15.104	10.035	8.280	6.407	5.686
		Lower Strong Bounds %					
E	48.267	48.267					
90–99	18.911	14.143	4.7677				
80–89	10.809	4.7698	2.9946	3.0442			
70–79	8.460	2.7894	0.6193	2.3603	2.6912		
60–69	7.352	1.4226	0.8894	0.1916	2.8988	1.9493	
<60	6.202	0.4461	0.1587	0.0958	0.3784	1.626	3.497
Sum	100	71.838	9.430	5.692	5.968	3.575	3.498
		FDH Model %					
E	48.267	48.267					
90–99	18.911	16.619	2.292				
80–89	10.809	7.9835	1.9051	0.92			
70–79	8.460	6.0866	0.6758	0.772	0.9259		
60–69	7.352	5.6504	0.8345	0.2392	0.463	0.1647	
<60	6.202	2.4329	0.1587	0.926	1.134	0.6388	0.9115
Sum	100	87.039	5.866	2.857	2.523	0.804	0.912

Conceptually, T^{CRS} is simply the free disposable hull of the smallest convex cone that contains the data. Here, the *free disposable hull* of a set $S \subset R_+^m \times R_+^p$ is defined as

$$S^F = \left\{(X, Y) \in R_+^m \times R_+^p : X \geq X', Y \leq Y' \text{ for some } (X', Y') \in S\right\}. \tag{8}$$

Let

$$R_j = \{(X, Y) = \lambda_j(X_j, Y_j), \lambda_j \geq 0\} \tag{9}$$

denote a ray emanating from the origin that passes through one of the observed data point (X_j, Y_j). It may be readily verified that equation (7) is equivalent to

$$T^{CRS} = \left(\sum_j R_j\right)^F. \tag{10}$$

Representation (10) shows that points in the CRS technology are generated by all points obtained as a sum of operations, each of whom represents a scaled version of an observed operation (DMU). Points in the FC technology are also generated by all points obtained as a sum of operations; however, each operation represents an *appropriately* scaled version of an observed operation. That is, the FC technology eliminates all points in R_j whose λ_j is positive but too small. Let

$$\Lambda_j = \{0\} \cup [\ell_j, \infty) \tag{11}$$

and define

$$T_j = \{(X, Y) = \lambda_j(X_j, Y_j), \lambda_j \in \Lambda_j\}. \tag{12}$$

The FC technology is defined as

$$T^{FC} = \left(\sum_j T_j\right)^F, \tag{13}$$

which is algebraically equivalent to

$$T^{FC} = \left\{(X, Y) \in R_+^m \times R_+^p : X \geq \sum_j \lambda_j X_j, Y \leq \sum_j \lambda_j Y_j \text{ for some } \lambda \in \prod_j \Lambda_j\right\}. \tag{14}$$

It is understood that $\ell_j \leq 1$ so that the empirical production possibility set T^{FC} contains the observed data.

To determine the efficiency score for the FC technology one simply adds the following integer constraints to the corresponding linear program (1–3):

$$\ell_j \cdot z_j \leq \lambda_j \leq M \cdot z_j, \quad j = 1, \ldots, N \tag{15}$$

$$z_j \in \{0, 1\}, \quad j = 1, \ldots, N. \tag{16}$$

When $z_j = 0$, the corresponding intensity must be zero; otherwise, when $z_j = 1$, the intensity must be larger than the lower bound ℓ_j. M is chosen sufficiently large or can be used to impose upper bounds to prevent cases as reported in Thore (1996). In this case, Λ_j in equation (11) should be changed accordingly.

Geometrically, the FC technology defines an efficiency frontier that lies between the two extreme frontier representations (CRS and FDH). Figure 1 depicts a 2-dimensional representation of the FC technology vis-a-vis some of the other known frontiers.

The CRS frontier (from $X_2 \to \infty$ to A, D, H and $X_1 \to \infty$) is the most "conservative" frontier in Figure 1. It is followed by the FC frontier. Both the CRS and the FC frontiers are depicted in the main body of the figure, as well as in the detailed view. The broken line connecting points A and D represents all possible convex combinations of these two points. However, only the line segment B–C is FC feasible and therefore the FC frontier between A and B is given by the line that starts at A and goes parallel to the X_1 axis till X_1^B and then parallel to the X_2 axis till X_2^B. To the right of C, the FC frontier goes parallel to X_1 axis till X_1^D and down, parallel to X_2 axis, till X_2^D. Similar explanation holds for the FC frontier segments that are on the right side of D. The next frontier is the NIRS model given by Petersen (1990) (ibid., Fig. 1, p. 309). To construct the NIRS frontier, we add points

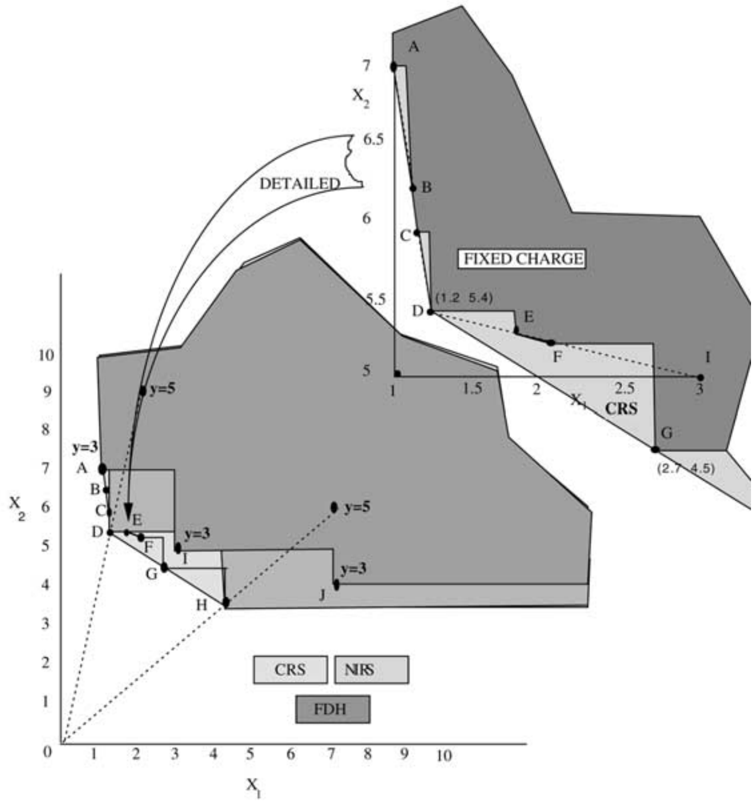

Figure 1. Fixed charge technology compared to known technologies.

D and H, which are scaled down (by 3/5) representations of the given data, and then we connect points A, D, H using the free-disposability principle. The most "liberal" frontier in Figure 1 is the FDH which goes from A parallel to X_1 till X_1^I, then parallel to X_2 till X_2^I and continues in the same manner to J.

The choice of lower bounds may be obtained from specific knowledge of each DMU's operations. Lower bounds may also be set to satisfy managerial perceptions about what constitutes an appropriate mix; that is, no DMU may appear in a composite unless it contributes at least a minimal percentage, etc. In what follows, for purposes of conducting empirical comparisons, we analyzed 3 choices for the lower bounds ℓ_j—see equation (5):

- *Strong restriction* $\ell_j = \min\{\ell_j^{in}, \ell_j^{out}\}$
- *Strong input restriction* $\ell_j = \ell_j^{in}$
- *Weak restriction* $\ell_j = \max_{i,r}\left\{\dfrac{X_i^{Min}}{X_{ij}}, \dfrac{Y_r^{Min}}{Y_{rj}}\right\}$

These restrictions are only well-defined if all inputs and outputs are strictly positive, which holds true in all of the datasets analyzed. Using the strong, strong input or weak restrictions ensure that no violation of the corresponding type will occur. Notice that the lower bounds required to guard against a strong violation are actually *lower* then the lower bounds required to guard against a weak violation. To avoid a strong violation only one input or output of the composite DMU must lie above the corresponding value of Min_DMU, whereas to avoid a weak violation *all* inputs and outputs must have this property. Since constraints are added to the standard minimization program, the efficiency scores of the new model cannot decrease.

Each of the 18 datasets was run using the new model. Results are summarized in Table 3. There are significant increases in the efficiency scores of the new models as compared with their original counterparts. There is an insignificant increase in the number of efficient DMUs when the strong restriction lower bounds are applied, whereas there is a significant increase in the number of efficient DMUs when the weak restriction lower bounds are applied.

4. Analysis and Geometry of CRS_FC

In this section we will confine our attention to the analysis and geometry of the CRS_FC model. The proposed DEA_FC models exhibit nonconvexities due to the fixed-charge constraint, which leads to fundamental differences between the new models and the standard ones. When convexity is imposed, the set of efficient points is sufficient to generate the technology. In sharp contrast to the basic CRS or VRS models, for the DEA-FC model *inefficient DMUs* may play a pivotal role in determining the technology. For example, *component DMUs may be inefficient and scaled versions of inefficient DMUS may be efficient.* We begin by explaining why low intensities exist in DEA.

4.1. The Low Intensity Phenomenon

To explain this phenomenon we need to develop some mathematical properties of partial efficiency scores. Let $\theta_k^*(S_i, N_j)$ denote the optimum solution of the linear program $P(S_i, N_j)$ defined by:

$$P(S_i, N_j): \quad \min \theta_k$$
$$\text{subject to}$$
$$\sum_{s \in N_j} \lambda_s X_s \leq \theta_k X_k, \tag{17}$$
$$\sum_{s \in N_j} \lambda_s y_{s\ell} \geq y_{k\ell}, \quad \ell \in S_i.$$

Here, the output set is restricted to set $S_i \subset \{1, 2, \ldots, p\}$ and the set of DMUs is restricted to set $N_j \subset \{1, 2, \ldots, N\}$.

PROPOSITION 1 *Let $S_1, \ldots, S_K$ be a partition of the set of outputs; let $N_1, N_2, \ldots, N_L$ be subsets of the set of DMUs, and let θ_k^* denote the optimum solution to the original CRS linear program. Then $\sum_{i,j} \theta_k^*(S_i, N_j) \geq \theta_k^*$.*

Table 4. Normalized data for the example.

	Normalized Inputs		Normalized Outputs			
	$\bar{I}_1$	$\bar{I}_2$	$\bar{O}_1$	$\bar{O}_2$	$\bar{O}_3$	$\bar{O}_4$
DMU_0	1.00	1.00	1.00	1.00	1.00	1.00
DMU_1	**0.09**	0.04	0.03	0.87	0.82	**0.68**
DMU_2	**1.27**	0.53	**51.60**	2.50	0.36	75.40

Proof. Let $\lambda^*(S_i, N_j)$ denote an optimal solution to $P(S_i, N_j)$. Extend $\lambda^*(S_i, N_j)$ to $\Re^{N+1}$ by adding the requisite zeroes in the obvious way—let $\hat{\lambda}(S_i, N_j)$ denote the extended vector. Define $\tilde{\lambda} = \sum_{i,j} \hat{\lambda}(S_i, N_j)$, and $\tilde{\theta}_k = \sum_{i,j} \theta_k^*(S_i, N_j)$. The result follows from the fact that $(\tilde{\lambda}, \tilde{\theta}_k)$ is a feasible solution for the CRS problem. ■

We will use this result in the following example. Table 4 lists the normalized inputs and outputs of 3 hypothetical DMUs. We chose DMU_0 as the reference unit and used its input–output values to normalize the data. Using the CCR model, the composite unit is composed by 0.02 of DMU_2 and 1.21 of DMU_1. The small intensity associated with DMU_2 is the result of a *scale effect* and *complement effect*. If DMU_0 is compared separately to DMU_1 or to DMU_2, it will be assessed as efficient. Stated differently, θ_0^* $(\{1, 2, 3, 4\}, \{j\}) = 100\%, j = 1, 2$. (Its FDH efficiency score is 1.) However, when both DMU_1 and DMU_2 are considered, DMU_0 is only 13% efficient. Why does this happen? The main problem is that DMU_1 produces very little output 1. If output 1 were not in the model so that $S_2 = \{2, 3, 4\}$, and if DMU_0 is compared only to DMU_1 so that $N_1 = \{1\}$, then the efficiency score becomes $\theta_0^*(S_2, N_1) = 100 \times (0.09/0.68) = 13.2\%$. On the other hand, DMU_2 produces a disproportional amount of output 1 per unit of input. So, if $S_1 = \{1\}$ and $N_2 = \{2\}$, then the efficiency score becomes $\theta_0^*(S_1, N_2) = 100 \times (1.27/51.6) = 2.5\%$ and not 100% as obtained by the FDH model. DMU_2 complements DMU_1, and since DMU_2 is exceptionally productive with respect to output 1, DMU_1 can afford to "buy" only a small portion.

In the above example DMU_1 can reveal DMU_0 to be extremely inefficient, if only S_1 is considered; similarly, DMU_2 can reveal DMU_0 to be extremely inefficient, if only S_2, the complementary set of outputs, is considered. To reveal DMU_0 efficient *all* outputs must be considered. It turns out that when all outputs are considered DMU_0 is still extremely inefficient. As Proposition 1 shows, this result is not an artifact of the particular numbers chosen: the inefficiencies of DMU_0 computed when the output set is restricted to two complementary subsets can be used to bound the inefficiency of DMU_0 when all outputs are considered, namely, $\hat{\theta}_0 = \theta_0(S_2, N_1) + \theta_0(S_1, N_2) = 15.7\% > 13\% = \theta_0^*$.

4.2. *Inefficient Component Decision-Making Units*

We begin by establishing that an inefficient DMU cannot be a component in its composite DMU, which holds true in the standard DEA models.

PROPOSITION 2 *An inefficient DMU cannot be a component in its composite DMU, i.e., if $\theta_0^* < 1$ then $\lambda_0^* = 0$.*

Proof. Suppose there exists a feasible solution (λ, θ) for which $\lambda_0 > 0$ and $\theta_0 < 1$. We shall demonstrate that (λ, θ) cannot be optimal by exhibiting a feasible solution $(\hat{\lambda}, \hat{\theta}_0)$ for which $\hat{\theta}_0 < \theta_0$, as follows: $\hat{\lambda}_0 = 0$, $\hat{\lambda}_j = \frac{\lambda_j}{1-\lambda_0}$, $j \neq 0$, and $\hat{\theta}_0 = \frac{\theta_0 - \lambda_0}{1-\lambda_0}$. It is easily verified that $(\hat{\lambda}, \hat{\theta}_0)$ is feasible, and that $\hat{\theta}_0 < \theta_0$, since $f(x) = \frac{\theta_0 - x}{1-x}$ is a decreasing function of $x \in [0, 1)$. ■

Next we turn to demonstrating that an inefficient DMU could belong to a reference set. Consider the following 2-input, 1-output example. There are 3 DMUs: DMU_A uses 8 units of labor and 1 unit of capital to produce 1 unit of output; DMU_B uses 1 unit of labor and 8 units of capital to produce 1 unit of output; and the reference DMU, DMU_0, uses 7 units each of labor and capital to produce 1 unit of output. Under CRS the reference DMU has an efficiency score of $4.5/7 = 64\%$—the composite DMU's components consist of 50% of each DMU_A and DMU_B.

Now suppose it is known that the engineering blueprint that defines the operations of DMU_A is such that a minimum of 1 unit of labor at all times is required to produce any output, and that the engineering blueprint that defines the operations of DMU_B is such that a minimum of 1 unit of capital at all times is required to produce any output. Quite often a unit of resource may have to be consumed even if its utilization rate is less than 100%. With these engineering requirements a lower bound of 1 applies to each of these DMUs. The unit level set, which defines the collection of all points in R^2 can achieve unit output, is depicted in Figure 2. (Note that a lower bound of 1 also ensures that no weak violation will occur since in this example Min_DMU uses 1 unit of labor and capital.) When the lower bounds are applied, DMU_0 is easily seen to be input efficient. Note that by combining the

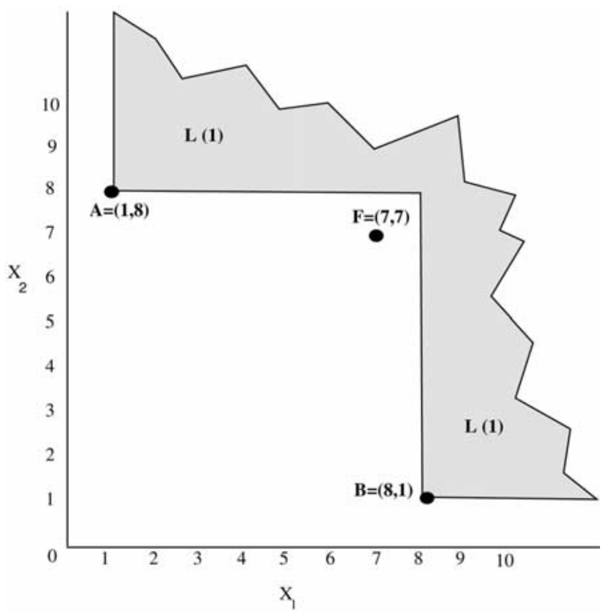

Figure 2. $L(1)$ generated by DMUs A, B.

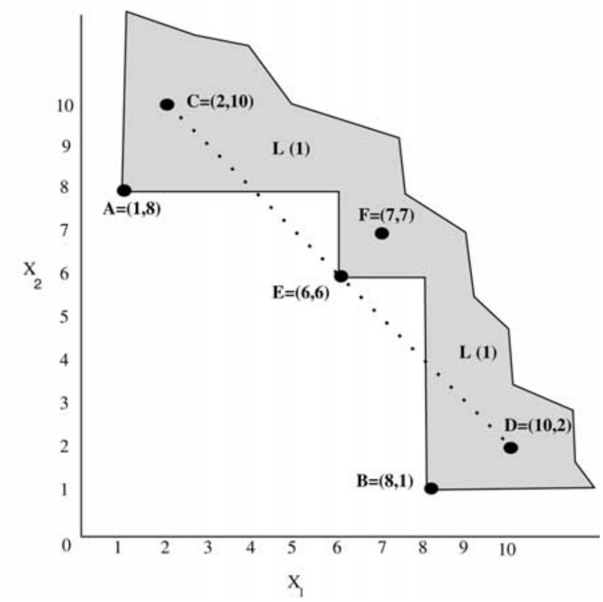

Figure 3. $L(1)$ generated by DMUs A, B, C, D.

operations of both DMUs at full scale, 9 units each of labor and capital can produce 2 units of output. Thus, a 30% increase in input would double the output.

Now suppose two new DMUs are added to the dataset, as follows. DMU_C uses 10 units of labor and 2 units of capital to produce 1 unit of output, and DMU_D uses 2 units of labor and 10 units of capital to produce 1 unit of output. Each operation may be scaled as long as there is at least a unit of labor and capital, and so the lower bounds for these new DMUs are set to 50%. It is clear that DMU_C and DMU_D are each input inefficient. However, a 50–50% mixture of these DMUs determines a composite DMU_E that uses 6 units each of labor and capital to produce a unit of output. The new unit level set is depicted in Figure 3, which reveals DMU_0 to be input inefficient with an efficiency of $6/7 = 86\%$.

The inefficient DMUs, C and D, are thus required to determine the correct unit level set. But there is more. Points C and D may each be scaled by 50% to produce points $F = (1, 5)$ and $G = (5, 1)$, each of whom has output 0.5. The level sets corresponding to output production of 0.5 and 1.0 are depicted in Figure 4. Thus, *scaled* versions of the inefficient DMUs, C and D, are actually efficient.

4.3. Geometry of CRS_FC

We now turn to describing the geometry of CRS_FC, which is fundamentally different than the CRS or FDH models. It will be discussed for the case of 2 inputs and scalar output.

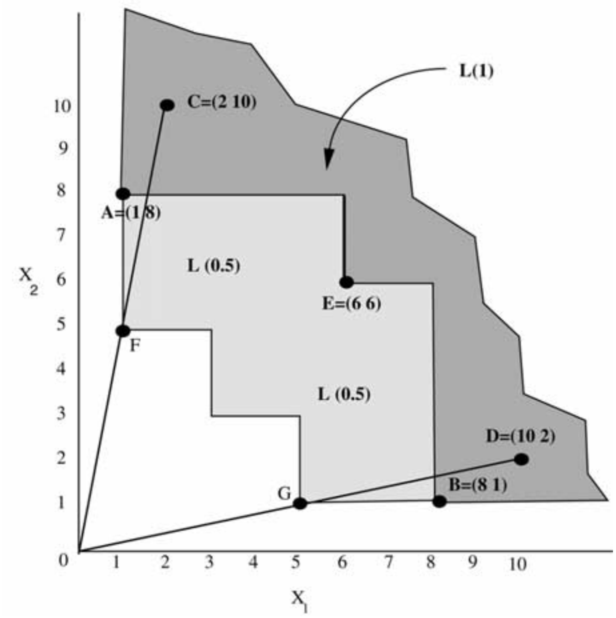

Figure 4. $L(1)$ and $L(0.5)$

The level set $L^{FC}(y)$ for CRS_FC is defined as follows:

$$L^{FC}(y) = \left\{ X \in \Re_+^m : \exists \lambda : \sum_{j=0}^{N} \lambda_j X_j \leq X, \sum_{j=0}^{N} \lambda_j y_j \geq y, \lambda_j \in \{0\} \cup [l_j, \infty) \right\}.$$

For each nonempty subset $K \subset N$ define the *level-subset* $L_k^{FC}(y)$ as follows:

$$L_K^{FC}(y) = \left\{ X \in \Re_+^m : \exists \lambda : \sum_{j \in K} \lambda_j X_j \leq X, \sum_{j \in K} \lambda_j y_j \geq y, \lambda_j \geq l_j, j \in K \right\}. \quad (18)$$

To show how to geometrically construct $L^{FC}(y)$ it is sufficient to show how to geometrically construct each $L_K^{FC}(y)$, since $L^{FC}(y) = \cup_{K \subset N} L_K^{FC}(y)$. Letting $\hat{X}_K(l) \equiv \sum_{j \in K} l_j X_j$, $\hat{y}_K(l) \equiv \sum_{j \in K} l_j y_j$, $\tilde{X} = X - \hat{X}_K(l)$ and $\hat{\lambda}_j = \lambda_j - l_j$, the level-subset $L_K^{FC}(y)$ is equivalent to:

$$L_K^{FC}(y) = \hat{X}_K(l) + \left\{ \begin{array}{ll} \tilde{X} \in \Re_+^m : \exists \hat{\lambda} : & \sum_{j \in K} \hat{\lambda}_j X_j \leq \tilde{X}, \\ & \sum_{j \in K} \hat{\lambda}_j y_j \geq y - \hat{y}_K(l), \\ & \hat{\lambda}_j \geq 0, j \in K \end{array} \right\}$$

The set defined in brackets is precisely the definition of the level set corresponding to output level $y - \hat{y}_K(l)$ of the CRS technology restricted to DMU dataset K. We shall denote

Table 5. Data for geometrical constructions.

DMU	Input 1	Input 2	Output
DMU_0	3	2.5	20
DMU_1	2	3	40
DMU_2	3	1	30

this set as $L_K^{CRS}(y - \hat{y}_K(l))$, and it is understood that this set coincides with $\Re_+^m$ when $y \leq \hat{y}_K(l)$. Thus,

$$L_K^{FC}(y) = \hat{X}_K(l) + L_K^{CRS}(y - \hat{y}_K(l)). \tag{19}$$

Recall that to construct $L_K^{CRS}(u)$ for any output $u > 0$ one calculates the *generator* points $X_j^* = (u/y_j)X_j$, for $j \in K$, forms their convex hull, and extends it by adding the nonnegative orthant to each X_j^*. Here, the generator points are simply

$$s_j^*(y) = \begin{cases} \hat{X}_K(l), & y \leq \hat{y}_K(l), \\ \hat{X}_K(l) + \frac{y - \hat{y}_K(l)}{y_j} X_j, & y > \hat{y}_K(l). \end{cases} \tag{20}$$

We now illustrate the level set construction and inherent nonconvexities of CRS_FC with an example. Table 5 lists the input and output of 3 DMUs: DMU_1, DMU_2 and the reference DMU_0.

The efficiency of DMU_0 is determined geometrically by intersecting the line that passes through the origin and the input vector of DMU_0 with the isoquant of the level set $L^{FC}(20)$.

Figure 5 depicts $L^{FC}(20)$ when the lower bounds are set to 0.00, 0.25, 0.33, 0.60, and 0.75, respectively. (For simplicity the lower bounds are identical for each DMU.) The calculations and algorithm necessary for the construction of each $L^{FC}(20)$ are given by Vlatsa (1995), while Table 6 summarizes the essential data. The graphs demonstrate the nonconvexity of the level set—the left-most and right-most portions of a line segment joining 2 boundary points will no longer belong to the level set, the degree of which depends entirely on the size of the lower bounds. It is also clear that $L^{FC}(20)$ moves to the upper right as the lower bounds increase until $\ell = 1.00$. The nonconvex level sets depicted herein fall within the class of projectively-convex sets, introduced in Hackman and Passy (1988) and further analyzed in First et al. (1993).

In all the cases depicted in Figure 5 both DMU_1 and DMU_2 were efficient. The outputs of the three DMU's are $y_0 = 20$, $y_1 = 40$ and $y_2 = 30$

(a)	$\ell = 0.00,$	$\theta_0^* = 0.46,$	$\lambda_1^* = 0.30,$	$\lambda_2^* = 0.27$
(b)	$\ell = 0.25,$	$\theta_0^* = 0.46,$	$\lambda_1^* = 0.30,$	$\lambda_2^* = 0.27$
(c)	$\ell = 0.33,$	$\theta_0^* = 0.55,$	$\lambda_1^* = 0.33,$	$\lambda_2^* = 0.33$
(d)	$\ell = 0.60,$	$\theta_0^* = 0.66,$	$\lambda_1^* = 0.00,$	$\lambda_2^* = 0.67$
(e)	$\ell = 0.75,$	$\theta_0^* = 0.75,$	$\lambda_1^* = 0.00,$	$\lambda_2^* = 0.75$

Table 6. Construction of L^{FC} (20) for different lower bound values.

		$\hat{X}_K(l)$			$X^*(l, K)$		s_i^*	
i	Subset K	$\hat{X}_1$	$\hat{X}_2$	$\hat{y}_K(l)$	x_1^*	x_2^*	s_{i1}^*	s_{i2}^*
$\ell = 0.25$								
(1)	$\{DMU_1\}$	0.50	0.75	10.00	0.50	0.75	1.00	1.50
(2)	$\{DMU_2\}$	0.75	0.25	7.50	1.25	0.42	2.00	0.67
(3)	$\{DMU_1, DMU_2\}$	1.25	1.00	17.50	0.13	0.19	1.38	1.19
(4)					0.25	0.08	1.50	1.08
(5)	$\{DMU_0\}$	0.75	0.63	5.00	2.25	1.88	3.00	2.50
$\ell = 0.33$								
(1)	$\{DMU_1\}$	0.67	1.00	13.33	0.33	0.50	1.00	1.50
(2)	$\{DMU_2\}$	1.00	0.33	10.00	1.00	0.33	2.00	0.66
(3)	$\{DMU_1, DMU_2\}$	1.67	1.33	23.33	—	—	1.67	1.33
(4)	$\{DMU_0\}$	1.00	0.83	6.66	2.00	1.67	3.00	2.50
$\ell = 0.6$								
(1)	$\{DMU_1\}$	1.20	1.80	24.00	—	—	1.20	1.80
(2)	$\{DMU_2\}$	1.80	0.60	18.00	0.20	0.07	2.00	0.67
(3)	$\{DMU_1, DMU_2\}$	3.00	2.40	42.00	—	—	3.00	2.40
(4)	$\{DMU_0\}$	1.80	1.50	12.00	1.20	1.00	3.00	2.50
$\ell = 0.7$								
(1)	$\{DMU_1\}$	1.50	2.25	30.00	—	—	1.50	2.25
(2)	$\{DMU_2\}$	2.25	0.75	22.50	—	—	2.25	0.75
(3)	$\{DMU_1, DMU_2\}$	3.75	3.00	52.50	—	—	3.75	3.00
(4)	$\{DMU_0\}$	2.25	1.88	15.00	0.75	0.63	3.00	2.50

5. Relaxation of Lower Bound Constraints: A Goal Programming Approach

Consider the following situation (adapted from the numerical analysis to follow). The VRS model was applied to a certain dataset and revealed DMU_0 as 61% efficient. The composite DMU exhibits a strong input violation. The position is taken that the composite DMU is not valid, and so the requisite lower bound constraints are added to ensure that DMU_0 will be more appropriately assessed. By doing so, DMU_0 is now considered efficient. Now suppose that another composite DMU can be found that only slightly violates the lower bound constraint (say by less than 20%), but renders a completely different verdict on DMU_0, namely, that it is now 90% efficient. This example motivates the following question: *Which assessment of efficiency is preferred?* On one hand, the efficiency rating should be fair—the scaled DMUs comprising the composite DMU should not be unreasonably low; on the other hand, this requirement should not render the discriminating power of DEA useless. There is a trade-off between what we shall term the *efficiency gap,* the difference between 0.61 and 1.00, and what we shall term the *deviation gap,* an indication of how far the solution is from the lower bound constraint.

To provide the decision-maker a quantitative assessment of the efficiency gap versus, the deviation gap we formulate the following class of Min-Max optimization

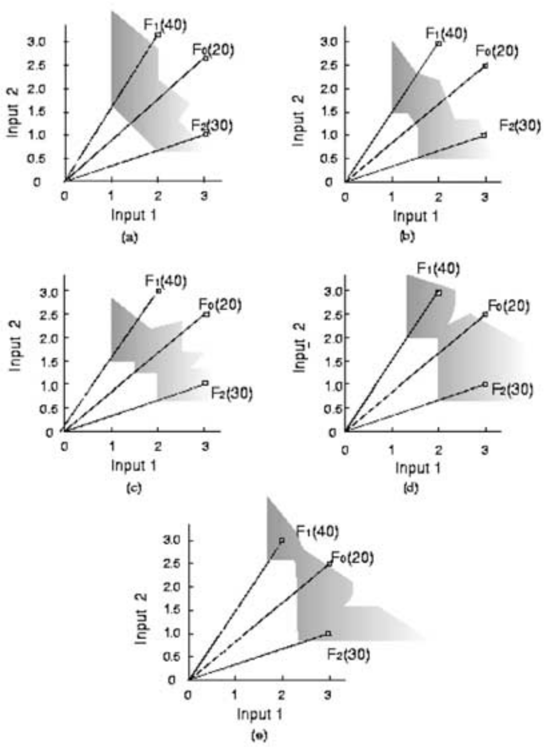

Figure 5. The CRS_FC dataset, L^{FC} (20), for different lower bounds.

problems:

$$
\begin{aligned}
\text{Min} \quad & f(q_1, \ldots, q_m) \\
\text{s.t.} \quad & \\
& q_i \geq \frac{\left(z_j X_i^{Min} - X_{ij}\lambda_j\right)}{W_i} \quad \forall i, j \\
& \lambda_j \leq z_j \quad \forall j \\
& \sum_j X_{ij}\lambda_j \leq \theta_0 X_{io} \quad \forall i \\
& \sum_j Y_{rj}\lambda_j \geq Y_{ro} \quad \forall r \\
& \sum_j \lambda_j = 1 \\
& \theta_0 \geq \theta_0^{LB} - \delta \\
& \lambda_j \geq 0 \quad \forall j;\, z_j \in \{0, 1\} \\
& q_i \geq 0 \quad \forall i
\end{aligned}
\tag{21}
$$

The parameter δ represents the permissible decrease in efficiency score from the efficiency score θ_0^{LB} that is obtained when the lower bound constraints are added to the VRS model. Since deviations from Min_DMU should only be counted for those DMUs that form the components of the composite DMU (i.e., for which $\lambda_j > 0$) the binary variable z_j is used to indicate association with the composite DMU (i.e., $z_j = 1$ if and only if $\lambda_j > 0$). Each q_i measures the maximum, normalized deviation of the difference between each component's use of input i and Min_DMU's use of input i. The W_i constants are used to normalize the deviation for each input i. We suggest that W_i is set to X_i^{Min}. With respect to measuring the deviation we suggest that the maximum deviation $f(q) = \max_i q_i$ is most appropriate.

To illustrate the trade-off between the efficiency gap and the deviation gap, we return to the 6 inefficient DMUs shown in Table 1. These trade-offs are graphically depicted in Figure 6. We focus first on DMU_6 which is the most inefficient DMU in this dataset. The efficiency gap ranges from 0.61, the VRS score with no lower bounds, to 1.00, when the strong input restriction constraints are added. When the efficiency is set to 0.61, the corresponding deviation is 0.60, which has the following meaning. First, each component of the composite DMU that is used to assess DMU_6 as 0.61 efficient violates the strong input restriction lower bound constraint by no more than 60%. That is, each component uses *no*

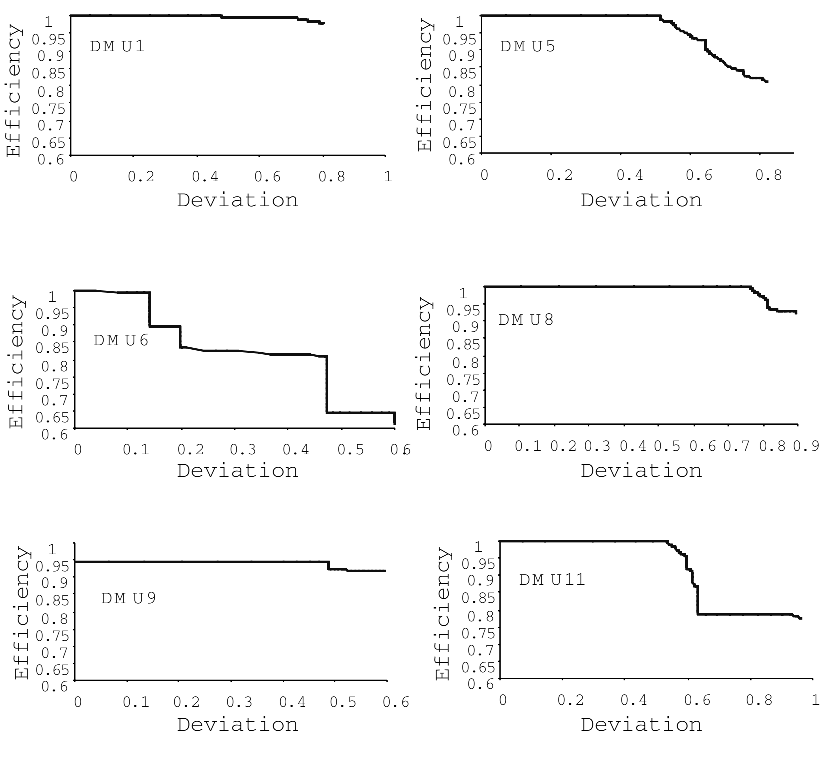

Figure 6. Efficiency vs. the deviation gap.

less than 40% of the inputs corresponding to Min_DMU with at least one input being used at exactly 40%. Second, each component of each composite DMU that may be used to assess DMU_6 as 0.61 efficient will have a strong input violation of *at least* 60%. The graph demonstrates that in order to assess DMU_6 as inefficient a sizeable relaxation of the lower bound constraint must be allowed. For example, an allowable deviation from the strong input restriction lower bound of say 33% corresponds to an efficiency of 0.82, approximately half of the efficiency gap.

The graphs of the efficiency versus the deviation gap for the other five inefficient DMUs tell a completely different story than the one for DMU_6. The efficiency gaps are significant for DMU_5 and DMU_{11}, whereas they are quite small for DMU_1, DMU_9 and DMU_8. A significant efficiency gap means that the choice of model, VRS versus. VRS_FC, will make a considerable difference in the assessment of efficiency. However, in all five cases, allowing a major violation of the lower bound constraints hardly makes a dent in the assessment of efficiency that was obtained using VRS_FC. For this dataset we can draw two conclusions about using the VRS_FC model: the assessments of efficiency (i) substantially differ from the VRS model; and (ii) they are extremely insensitive to perturbations of the lower bound constraints, i.e., they are *robust*.

Several potential extensions of the basic FC models should be noted. First, upper bounds could easily be added. Second, in lieu of a single Min_DMU, there could be one for each group of DMUs with the lower bounds set accordingly. For example, in the warehouse dataset groups could be classified according to size of facility (small, medium or large) or level of investment (low, medium or high). Third, with respect to the goal programming models, the average value $\bar{X}_i$ and the difference $\bar{X}_i - X_i^{Min}$ between the average and minimum values could be suitable choices for the normalization W_i. Fourth, instead of using the maximum deviation across all inputs the average deviation $f(q) = 1/m \sum_i q_i$ may be suitable.

6. Concluding Remarks

This paper extends the applicability of DEA by proposing a new family of models, which are formulated as mixed-integer mathematical programs. Conceptually, the new formulations are useful to model situations that exist in-between the CCR or BCC models and the FDH model. The focus of this paper has been on the possible introduction of lower bounds on intensities in the envelopment models. Obviously, we have not exhausted all such possible extensions. Topics of further research may include issues pertaining to upper bounds or other types of "fixed-charge" formulations, as well as the sensitivity of the new formulations to data perturbations. We also hope that our description and discussion of the nonconvexities associated with the new model in Section 4 leads to a sharper understanding of empirical production possibility sets and stimulates further inquiry into the implications of various modeling assumptions on the descriptions of technology.

Acknowledgment

We are grateful to the referees for comments that stimulated us to rethink the motivation, justification and implications for our model.

References

Baker, R. C. and S. Talluri. (1997). "A Closer Look at the Use of Data Envelopment Analysis for Technology Selection." *Computers & Industrial Engineering* 32(1), 101–108.

Banker, R. D., A. Charnes and W. W. Cooper. (1984). "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis." *Management Science* 30(9), 1078–1092.

Banker, R. D. and R. C. Morey. (1986). "The Use of Categorical Variables in Data Envelopment Analysis." *Management Science* 32(12), 1613– 1627.

Byrnes, P., R. Färe and S. Grosskopf. (1984). "Measuring Productive Efficiency: An Application to Illinois Strip Mines." *Management Science* 30(6), 671–681.

Charnes, A. and W. W. Cooper. (1980). "Auditing and Accounting for Program Efficiency in Non-for-Profit Entities." *Accounting Organizations and Society* 5(1), 87–107.

Charnes, A., W. W. Cooper, B. Golany, L. M. Seiford and J. Stutz. (1985). "Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions." *Journal of Econometrics* 30, 91–107.

Charnes, A., W. W. Cooper and S. Li. (1989). "Using Data Envelopment Analysis to Evaluate Efficiency in Economic Performance of Chinese Cities." *Socio-Economic Planning Sciences* 23(6), 325–344.

Charnes, A., W. W. Cooper and E. Rhodes. (1978). "Measuring the Efficiency of Decision Making Units." *European Journal of Operational Research,* 429–444.

Charnes, A., W. W. Cooper, A. R. Lewin and L. M. Seiford. (1995). *Data Envelopment Analysis: Theory, Methodology and Applications.* Kluwer Academic Publishers.

Charnes, A., W. W. Cooper, Q. L. Wei and Z. M. Huang. (1989). "Cone Ratio Data Envelopment Analysis and Multi-Objective Programming." *International Journal of Systems Sciences* 20, 1099–1118.

Chilingerian, J. A. (1995). "Exploring Why Some Physicians' Hospital Practices Are More Efficient: Taking the DEA Inside the Hospital." In A. Charnes, W. W. Cooper, A. R. Lewin and L. M. Seiford (eds.), *Data Envelopment Analysis: Theory, Methodology and Applications.* Kluwer Academic Publishers, 167–194.

Cook, W., Y. Roll and A. Kazakov. (1990). "A DEA Model for Measuring the Relative Efficiency of Highway Maintenance Patrols." *INFOR* 28(2), 113–124.

Crino, J. R., R. A. Maurer and F. Grange. (1996). "A DEA Evaluation of US Army Infantry Training Efficiency." Working Paper, Golden, CO: Colorado School of Mines.

Dyson, R. G. and E. Thanassoulis. (1988). "Reducing Weight Flexibility in Data Envelopment Analysis." *Journal Operational Research Society* 39(6), 563–576.

First, Z., S. T. Hackman and U. Passy. (1993). "Efficiency Estimation and Duality Theory for Non-Convex Technologies." *Journal of Mathematical Economics* 22, 295–307.

Frazelle, E. H. and S. T. Hackman. "The Warehouse Performance Index: A Single Point Metric for Benchmarking Warehouse Performance." Material Handling Research Center Technical Report TR-93-14, Georgia Institute of Technology, Atlanta, GA.

Giokas, D. (1991). "A Bank Branch and Bound Efficiency: A Comparative Application of DEA and the Log-linear Model." *Omega* 19(6), 549–557.

Golany, B. and S. Thore. (1997). "Restricted Best Practice Selection in DEA: An Overview with a Case Study Evaluating the Socio-Economic Performance of Nations." *Annals of Operations Research* 73, 117–140.

Hackman, S. T. and U. Passy. (1988). "Projectively-Convex Sets and Functions." *Journal of Mathematical Economics* 17, 51–60.

Hackman, S. T., E. H. Frazelle, P. M. Griffin, S. O. Griffin and D. A. Vlatsa. (2001). "Benchmarking Warehousing and Distribution Operations: An Input-Output Approach." *Journal of Productivity Analysis* 16(1), 79–100.

Petersen, N. C. (1990). "Data Envelopment Analysis on a Relaxed Set of Assumptions." *Management Science* 36(3), 305–314.

Roll, Y., W. Cook and B. Golany. (1991). "Factor Weights in Data Envelopment Analysis." *IIE Transactions* 23(1), 2–9.

Seiford, L. M. (1995). "A DEA Bibliography." In A. Charnes, W. W. Cooper, A. R. Lewin and L. M. Seiford (eds.), *Data Envelopment Analysis: Theory, Methodology and Applications.* Kluwer Academic Publishers, 437–471.

Seiford, L. M. and R. M. Thrall. (1990). "Recent Developments in DEA." *Journal of Econometrics* 46, 7–38.

Shephard, R. W. (1970). "Theory of Cost and Production Functions." Princeton, N.J: Princeton University Press.

Sinuany-Stern, Z., A. Mehrez and A. Barboy. (1994). "Academic Departments Efficiency via DEA." *Computers & Operations Research* 21(5), 543–556.

Sueyoshi, T. (1992). "Measuring the Industrial Performance of Chinese Cities by Data Envelopment Analysis." *Socio-Economic Planning Sciences* 26(2), 75–88.

Sueyoshi, T. (1997). "Measuring Efficiencies and Returns to Scale of Nippon Telegraph & Telephone in Production and Cost Analysis." *Management Science* 43(6), 779–798.

Thore, S. (1996). "Economies of Scale in the US Computer Industry: An Empirical Investigation Using Data Envelopment Analysis." *Journal of Evolutionary Economics* 6, 199–216.

Tulkens, H. (1993). "ON FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts and Urban Transit." *Journal of Productivity Analysis* 4(1/2), 183–210.

Vlatsa, D. A. (1995). "Data Envelopment Analysis with Intensity Restrictions." Unpublished Ph.D. Thesis, Georgia Institute of Technology, Atlanta, Georgia, USA, September.