Les Cahiers du GERAD ISSN: 0711–2440

On Characterizing Full Dimensional Weak Facets in DEA with Variable Returns to Scale Technology

M. Davtalab-Olyaie, I. Roshdi,
V. Partovi Nia, M. Asgharian

G–2013–97

December 2013

On Characterizing Full Dimensional Weak Facets in DEA with Variable Returns to Scale Technology

Mostafa Davtalab-Olyaie

Faculty of Mathematics and Computer Science
Kharazmi University, Tehran, Iran

m.davtalab@khu.ac.ir

Israfil Roshdi

Department of Mathematics, Science and Research Branch
Islamic Azad University, Tehran, Iran

i.roshdi@gmail.com

Vahid Partovi Nia

GERAD & Department of Mathematics and Industrial Engineering
Polytechnique Montréal, Montréal (Québec) Canada

vahid.partovi-nia@polymtl.ca

Masoud Asgharian

Department of Mathematics and Statistics
McGill University, Montréal (Québec) Canada

masoud@math.mcgill.ca

December 2013

Les Cahiers du GERAD

G–2013–97

Abstract: The frontier of the Production Possibility Set (PPS) consists of two types of full dimensional facets; efficient and weak facets. Identification of all facets of the PPS can be used in sensitivity and stability analysis, to find the closet target for an inefficient Decision Making Units (DMUs), and to determine the status of Returns to Scale of a DMU, among others. There has been a surge of articles on determining efficient facets in recent years. There are, however, many cases where knowledge of weak facets is required for a thorough analysis. This is the case, in particular, when the frontier of the PPS is constructed only of weak facets. The existing algorithms for finding weak facets either require knowledge of all extreme directions of the PPS or applicable only under some restrictions on the position of weak efficient DMUs. We provide a complete characterization of weak facets. Using this characterization we then devise a simple algorithm to find weak facets. Our algorithm can easily be implemented using existing packages for operation research, such as GAMS. We illustrate our algorithm using a numerical example.

Key Words: Data envelopment analysis; Full dimensional weak facet; Mixed integer linear programming; Polyhedral sets.

1 Introduction

Data envelopment analysis (DEA) is a mathematical programming method for evaluating the relative efficiency of Decision Making Units (DMUs) with multiple inputs and multiple outputs. The relative comparison in DEA is performed within the production possibility set (PPS) obtained from a set of postulates (see for example [1] and [2]). The performance level of each DMU is evaluated with respect to the frontier of the PPS. This performance level is known as the efficiency score, which only depends on the identified frontier. Identifying the frontier of the PPS is of prime interest in DEA. The frontier of the PPS comprises two types of facets: full dimensional efficient facets (FDEFs) and full dimensional weak facets (FDWFs).

Several authors have discussed FDEFs, e.g., [3, 4, 5, 6, 7] and [8]. There are, however, many cases where the PPS is composed only of weak facets. A trivial example is when the number of extreme efficient DMUs lying on each facet is less than the total number of inputs and outputs. Moreover, as discussed in [8], finding FDWFs can help us provide a more detailed analysis of data. In fact, having found all the full dimensional facets, i.e., both FDEFs and FDWFs, one can identify the reference set for an inefficient DMU [5], find the closest target for an inefficient DMU [9], group DMUs according to their frontiers [10], and use them for sensitivity and stability analysis [11, 12].

[13] have studied the structure and characteristics of weak surfaces of the PPS. Although their results are theoretically of interest, their algorithm requires knowledge of all extreme directions of the PPS. It is known that finding all extreme directions of a polyhedral set is a cumbersome task that can hinder applicability of their algorithm. [14] have provided an algorithm for finding the FDWFs of PPS with variable returns to scale technology. The algorithm devised by [14] ,as stated in their article, is not applicable if there are two DMUs that differ only in one element, output or input. Besides, one of their key results (Theorem 2 on page 3328) fails to hold.

In this article we take a different perspective. We show that the extreme (basic) optimal and non-positive solutions of both input and output oriented multiplier forms of the BCC model are the gradients of weak defining hyperplanes (WDHs), i.e., those hyperplanes whose intersection with the PPS produce the FDWFs, when evaluating a BCC-efficient DMU. To achieve our goal, we provide a detailed characterization of the structure of the FDWFs of the BCC-technology, using basic concepts of polyhedral set theory. We then use mixed integer linear programming (MILP) to devise an algorithm for determining the extreme optimal non-positive solutions of the input and output oriented multiplier forms of the BCC model, and hence find all FDWFs of the PPS.

The reminder of this manuscript is organized as follows. The input and output oriented primal and dual forms of the PPS are reported in Section 2 with a focus on the representation of the variable returns to scale technology. We further discuss the relationship between WDHs and extreme optimal solutions of input/output oriented BCC models in this section. We then study the structure of the FDWFs of the PPS with BCC technology in Section 2. Our algorithm is described in Section 3. We illustrate our algorithm in Section 4.

2 Structure of FDWFs

In this section, we will use some well-known concepts of polyhedral set theory, such as convex set, face, facet, extreme point, affinely independent set, recession direction etc., found in any standard book on optimization, see for example [15] or [16]. We first introduce the necessary notation and define the basic concepts of DEA used in this paper. Our main reference for these concepts is [17]. Throughout this paper, we deal with n observed homogeneous DMUs, each consuming m inputs to produce s outputs. The input and output vectors of DMU$_j$ for $j = 1, \dots, n$, where $x_j \neq 0$ and $y_j \neq 0$ are respectively denoted by $x_j = (x_{1j}, \dots, x_{mj})^T \in \mathbf{R}^m_{\geq 0}$ and $y_j = (y_{1j}, \dots, y_{sj})^T \in \mathbf{R}^s_{\geq 0}$. We further assume that o is the index of the under assessment DMU. After identifying inputs and outputs and gathering corresponding data in DEA, an appropriate production possibility set (PPS) will be determined. The PPS is the set of all technologically possible input-output combinations given by the following production technology:

$$T = \{(x, y) \mid \text{the output } y \text{ can be produced with the input } x\}. \tag{1}$$

Under the standard assumptions of inclusion of observations, convexity, variable returns to scale and free disposability of inputs and outputs, the unique non-empty PPS generated from n observed DMUs, DMU_j, $j = 1, \ldots, n$, can be represented in the following algebraic form:

$$T_v = \left\{ (x, y) \mid x \geq \sum_{j=1}^{n} \lambda_j x_j, y \leq \sum_{j=1}^{n} \lambda_j y_j, \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, j = 1, \ldots, n \right\}. \tag{2}$$

With reference to T_v, [2] introduced the following models for measuring the efficiency of DMU$_o$:

$$\begin{aligned} \theta_o = \min \quad & \theta - \epsilon\Big(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\Big) \\ \text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta x_{io}, i = 1, \ldots, m \\ & \sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro}, r = 1, \ldots, s \\ & \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, j = 1, \ldots, n, \\ & s_i^- \geq 0, i = 1, \ldots, m, s_r^+ \geq 0, r = 1, \ldots, s, \\ & \theta \text{ free.} \end{aligned} \tag{3}$$

$$\begin{aligned} \phi_o = \max \quad & \phi + \epsilon\Big(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\Big) \\ \text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io}, i = 1, \ldots, m \\ & \sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = \phi y_{ro}, r = 1, \ldots, s \\ & \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, j = 1, \ldots, n, \\ & s_i^- \geq 0, i = 1, \ldots, m, s_r^+ \geq 0, r = 1, \ldots, s, \\ & \theta \text{ free,} \end{aligned} \tag{4}$$

where ϵ is a non-Archimedean small and positive number and s_i^- and $s_r^+, i = 1, \ldots, m, r = 1, \ldots, s$, are non-negative slack variables. The variables s_i^- and s_r^+ represent input excesses and output shortfalls, respectively. Models 3 and 4 are called the input and output oriented envelopment forms, respectively. DMU_o is BCC-efficient or strong efficient if every optimal solution of model 3 (model 4), $\big(\theta^*, \lambda^*, s_i^{-*}, s_r^{+*}\big)$ $\big(\big(\phi^*, \lambda^*, s_i^{-*}, s_r^{+*}\big)\big)$, satisfies $\theta^* = 1$, $s_i^{-*} = 0, s_r^{+*} = 0, i = 1, \ldots, m, r = 1, \ldots, s$ ($\phi^* = 1$, $s_i^{-*} = 0, s_r^{+*} = 0, i = 1, \ldots, m, r = 1, \ldots, s$). Also, DMU_o is weak efficient if there exists an optimal solution of model 3 (model 4) such that $\theta^* = 1$ and $(s^{-*}, s^{+*}) \neq (0, 0)$ ($\phi^* = 1$ and $(s^{-*}, s^{+*}) \neq (0, 0)$). The BCC-efficient DMU_o is an extreme efficient DMU if the model 3 (model 4) has unique optimal solution with $\lambda_o^* = 1, \lambda_j^* = 0, \forall j \neq o$, otherwise is non-extreme efficient. To employ our algorithm, presented in Section 3, we need to determine all extreme BCC-efficient DMUs, say E. We can use approaches presented in [18] and [8] to obtain all extreme BCC-efficient DMUs.

The dual of models 3 and 4 without ϵ, i.e., with $\epsilon = 0$, are called the input and output oriented multiplier forms, denoted below by models 5 and 6, respectively.

$$\begin{aligned}\max\quad & U^t y_o + u_0 && (5)\\ \text{s.t.}\quad & V^t x_o = 1, \\ & U^t y_j - V^t x_j + u_0 \le 0, j = 1, \ldots, n, \\ & U \ge 0, V \ge 0.\end{aligned}$$

$$\begin{aligned}\max\quad & V^t x_o - u_0 && (6)\\ \text{s.t.}\quad & U^t y_o = 1, \\ & U^t y_j - V^t x_j + u_0 \le 0, j = 1, \ldots, n, \\ & U \ge 0, V \ge 0.\end{aligned}$$

DMU$_o$ is BCC-efficient, if there exists at least one optimal solution (U^*, V^*, u_0^*) of model 5 (model 6), with $(U^*, V^*) > 0$ and $U^{*t} y_o + u_0^* = 1$ $(V^{*t} x_o - u_0^* = 1)$; otherwise, DMU$_o$ is BCC-inefficient.

It is well known that if (U^*, V^*, u_0^*) is an optimal solution of model 5 (model 6), $H^* : U^{*t} y - V^{*t} x + u_0^* = 0$ is a supporting hyperplane of the T_v, i.e., the inequality $U^{*t} y - V^{*t} x + u_0^* \le 0$ is valid for all $(x, y) \in T_v$. Thus, the set

$$F = T_v \cap \left\{(x, y) \in R^{m+s} : U^{*t} y - V^{*t} x + u_0^* = 0\right\} = H^* \cap T_v$$

is a face of T_v. To completely characterize the structure of an FDWF of T_v, we need the following definitions and preliminaries.

Definition 1 *(weak and strong face). If in the optimal solution of model 5 or 6, (U^*, V^*, u_0^*), (U^*, V^*) is not strictly positive (positive), then $H^* = U^{*t} y - V^{*t} x + u_0^* = 0$ is called* weak (strong) supporting *and the corresponding face, $F^* = H^* \cap T_v$, is called* weak (strong) face.

Definition 2 *(full dimensional efficient facet, strong defining hyperplane). Suppose that $H_o : U^t y - V^t x + u_0 = 0$ is a supporting hyperplane of T_v. The set $F = H \cap T_v$ is called a* full dimensional efficient facet *(FDEF) of T_v and H is called* strong defining hyperplane *(SDH) if*

1. *(F is a strong face) all multipliers are strictly positive i.e. $(U, V) > 0$, and*
2. *(F is a facet of T_v) there exists at least one affine independent set with $m + s$ elements of BCC-efficient DMUs lying on F.*

Definition 3 *(full dimensional weak facet, weak defining hyperplane). Suppose that $H^* = U^{*t} y - V^{*t} x + u_0^* = 0$ is a supporting hyperplane of T_v. The set $F^* = H^* \cap T_v$ is called an* full dimensional weak efficient *(FDWF) of T_v and H^* is called* weak defining hyperplane *(WDH) if*

1. *(F^* is a weak face). There is at least one zero multiplier in variables U and V, and*
2. *(F^* is a facet of T_v) there exists at least one affine independent set with $m + s$ elements of BCC-efficient or weak efficient (virtual or observed) DMUs lying on F^*, i.e. the dimension of F^* is $m + s - 1$.*

To provide a visual representation of the above definitions, we give the following example. Consider DMU$_o$ in Figure 1. Using model 5, it can be seen that there are infinitely many supporting hyperplanes passing through DMU$_o$, corresponding to each optimal solution of model 5, of which only two hyperplanes, $H_1$and H_2, are defining and $F_1 = H_1 \cap T_v$ (weak facet), $F_2 = H_2 \cap T_v$(strong facet) are full dimensional facets of T_v.

[7] have proposed a method to determine all strong defining hyperplanes (SDHs) based on finding all the extreme positive solutions of model 5. Indeed, they have proved that there is a one-to-one correspondence between the extreme optimal solutions of model 5 in evaluating a BCC-efficient DMU and the gradients of SDHs of T_v. We show here that, similar to SDHs, such statement can be established for WDHs of T_v. In other words, we prove that there exists a one-to-one correspondence between the extreme and non-positive optimal solutions of model 5 (model 6) in evaluating DMU$_o$ and the gradient of WDHs of T_v with $V \neq 0$ $(U \neq 0)$ passing through DMU$_o$ in Theorems 2.1 and 2.2 below.

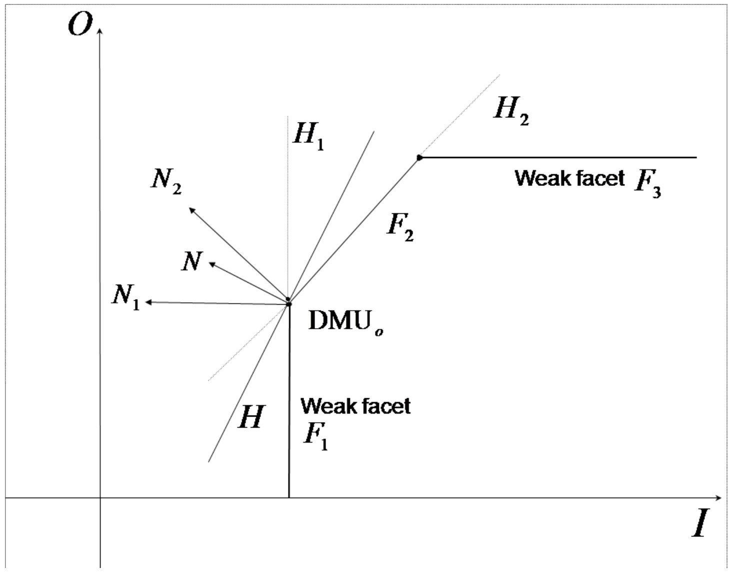

Figure 1: A PPS constructed by two DMUs, with one input, horizontal axis, and one output, vertical axis. Infinitely many supporting hyperplanes passing through DMU_o with different gradients are shown by N, N_1 and N_2. The hyperplanes H_1 and H_2 are defining for the PPS and H is supporting but not defining. The facets F_1 and F_3 of the PPS are weak, but F_2 is strong.

Theorem 2.1 *Suppose that (U^*, V^*, u_0^*) is an extreme optimal solution with some zero element in (U^*, V^*) when evaluating a BCC-efficient DMU, DMU_o, by model 5 (model 6), then $H_o^* : U^{*t}y - V^{*t} + u_0^* = 0$ is a WDH passing through DMU_o and therefore $F^* = H^* \cap T_v$ is an FDWF of T_v.*

See Appendix for proof.

It is worth mentioning that the inverse of the above theorem is also valid if $V^{*t}x_o \neq 0$ ($U^{*t}y_o \neq 0$). Having assumed that all inputs (outputs) are positive, we have the following theorem.

Theorem 2.2 *Suppose that $H_o^* : U^{*t}y - V^{*t}x + u_0^* = 0$ is a WDH passing through DMU_o in which $V^* \neq 0$ ($U^* \neq 0$). Then (U^*, V^*, u_0^*) is an extreme optimal solution of model 5 (model 6).*

See Appendix for proof.

We can therefore use models 5 and 6 for determining all WDHs of T_v passing through each extreme BCC-efficient DMU, and eventually determine all WDHs of T_v. To do so, we first find all WDHs with $V \neq 0$ by determining all extreme optimal non-positive solutions of (5) for each extreme BCC-efficient DMU. However, there may exist some WDHs with $V = 0$. For instance, consider DMU_o in Figure 2. As can be seen, the WDH $H_o : u = y_c$ in which $v_1 = v_2 = 0$, passes through DMU_o. Obviously, the gradients corresponding to these WDHs can not be a part of feasible solutions of model5 while they are the extreme optimal and non-positive solutions of model 6 as we have shown in Theorems 2.1 and 2.2. Therefore, for finding WDHs with $V = 0$, it suffices to determine all extreme optimal solutions of the following model which are obtained by setting $V = 0$ in model 6, for each extreme BCC-efficient DMU.

$$\begin{aligned} \max \quad & u_0 \\ \text{s.t.} \quad & U^t y_o = 1, \\ & U^t y_j \leq u_0, j = 1, \ldots, n, \\ & U \geq 0. \end{aligned} \tag{7}$$

It is worth mentioning that finding WDHs with $V = 0$ is equivalent to determining the gradients of all defining hyperplanes of PPS constructed by only the output components of DMU$_j$, $j = 1, \ldots, n$, i.e.,

$$\hat{T}_v = \left\{ y \mid y \leq \sum_{j=1}^{n} \lambda_j y_j, \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, j = 1, \ldots, n \right\}. \tag{8}$$

Indeed, if $\hat{U}$ is an extreme optimal solution of model 7, then $(U, V, u_0) = (\hat{U}, 0, -1)$ is an extreme optimal solution of model 6. Therefore $\hat{H} : \hat{U}^t y = 1$ is a WDH of T_v. Let

$$\hat{E} = \{j \mid y_j \text{ is extreme BCC-efficient of } \hat{T}_v\};$$

It is easy to see that $\hat{E} \subseteq E$. We then need only determine gradients of all defining hyperplanes passing through DMUs in $\hat{E}$. This, in turn, is equivalent to finding all extreme optimal solutions of model 7 for any members of $\hat{E}$. This follows from the above theorems and [7]. To summarize, we need to obtain all the extreme optimal and non-positive solutions of model 5 for each member of E and all the extreme optimal solutions of model 7 for any member of $\hat{E}$. The characterization of the structure of the FDWFs of the BCC-technology given in Theorem 2.3 below can help us to achieve this goal. Using Theorem 2.3 we devise an algorithm to find all FDWFs of the PPS.

It is clear that $\bar{F} = \bar{H} \cap T_v$ is a polyhedral where $\bar{H} = \bar{U}^t y + \bar{V}^t x + \bar{u}_0 = 0$ is a weak supporting hyperplane of T_v. In the classical representation theorem for the polyhedral sets, the extreme points and extreme directions of a polyhedral play a significant role in the parametric description of the polyhedral. Obviously, if DMU$_o$ is an extreme BCC-efficient lying on $\bar{F}$, it is an extreme point of the set $\bar{F}$. Theorem 2.3 below determines all types of extreme directions of $\bar{F}$.

Theorem 2.3 *Suppose that DMU$_o = (x_o, y_o)$ is an extreme BBC-efficient DMU and $\bar{H} = \bar{U}^t y + \bar{V}^t x + \bar{u}_0 = 0$ is a supporting hyperplane of T_v passing through DMU$_o$.*

1. *If DMU$_k = (x_k, y_k)$ is an extreme BBC-efficient DMU lying on $\bar{H}$, $d_k = (x_k - x_o, y_k - y_o)$ is an extreme feasible direction of the face $\bar{H} \cap T_v$ at point DMU$_o$.*
2. *If the tth element of the vector $\bar{V}$ is equal to zero, $\bar{v}_t = 0$, then $d_t = (e_t, 0)$ is an extreme recession direction of the face $\bar{H} \cap T_v$.*
3. *If the qth element of the vector $\bar{U}$ is equal zero, $\bar{u}_q = 0$, there are possibly two cases:*
   - *If $y_{qo} > 0$, then $d_q = (0, -e_q)$ is an extreme feasible direction of the face $\bar{H} \cap T_v$ at point DMU$_o$.*
   - *If $y_{qo} = 0$, then $d_q = (0, -e_q)$ is not a feasible direction of the face $\bar{H} \cap T_v$ at point DMU$_o$.*

See Appendix for proof.

Remark 1 *Let $D = \{d_1, \ldots, d_k\}$ be the set of all extreme feasible and recession directions of the face $\bar{F}$ at point DMU$_o$. If $\bar{H}$ is a strong face, i.e. $(\bar{U}, \bar{V}) > 0$, then the set D includes the only direction of type 1. If $\bar{H}$ is a weak face, the set D must include at least one direction of type 2 or 3 and perhaps some directions of type 1.*

Figure 2 depicts all three types of directions.

3 Identification of FDWFs

We first present a mixed binary program to determine the WDHs of T_v with $V \neq 0$; equivalently, for determining extreme optimal and non-positive solutions of model 5. Suppose that $H = U^t y - V^t x + u_0 = 0$ is a weak supporting hyperplane of T_v passing through DMU$_o$ with $V \neq 0$. Any optimal solution $(\hat{U}, \hat{V}, \hat{u}_0)$ of model 5 that fulfills the following two conditions is a gradient of a WDH with $V \neq 0$.

- $\hat{F} = \hat{H} \cap T_v$ where $\hat{H} = \hat{U}^t y - \hat{V}^t x + \hat{u}_0 = 0$ should be a proper face with the highest possible dimension, i.e. $\hat{F}$ should have maximum number of directions.

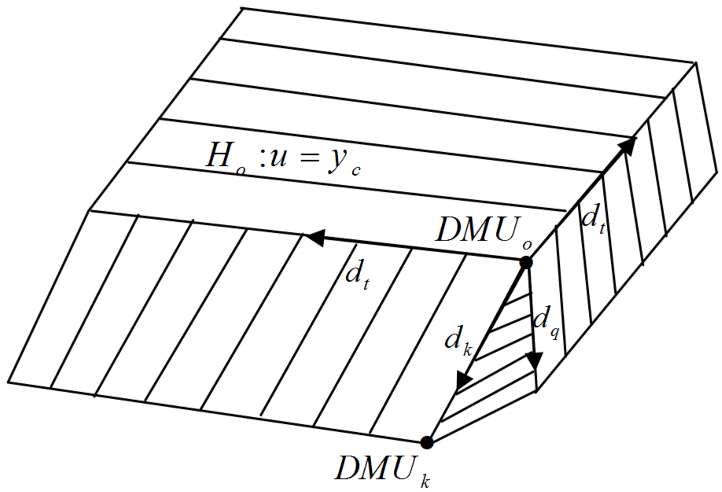

Figure 2: Various types of directions, d_k, d_t and d_q, at point DMU$_o$, for FDWFs of T_v.

- $\hat{H} = \hat{U}^t y - \hat{V}^t x + \hat{u}_0 = 0$ should be a weak supporting hyperplane of T_v, i.e. $\hat{F}$ should certainly have at least one direction of either type 2 or 3.

To guarantee that the first condition is met, we should choose an optimal solution of model 5 so that the total zero elements in the vector (U, V) and the number of extreme BCC-efficient DMUs lying on its corresponding hyperplane is maximized. As for the second condition, we add a constraint to our mixed integer linear program to control the total number of type 2 and 3 directions.

Using the above discussion, we can formulate the following mixed integer linear program to obtain an optimal solution of model 5 that corresponds to the gradient of a WDH of T_v. This result is formally established in Theorem 3.1.

$$
\begin{aligned}
\min \quad & I_o = \sum_{i=1}^{m} p_i + \sum_{r=1}^{s} q_r + \sum_{j \in E-\{o\}} I_j \qquad (9)\\
\text{s.t.} \quad & V^t x_o = 1,\\
& U^t y_o + u_0 = 1,\\
& U^t y_j - V^t x_j + u_0 + t_j = 0, j \in E - \{o\},\\
& t_j - M I_j \leq 0, j \in E, (*)\\
& v_i - M_1 p_i \leq 0, i = 1, \ldots, m, (**)\\
& u_r - M_2 q_r \leq 0, r = 1, \ldots, s, (***)\\
& \sum_{i=1}^{m} p_i + \sum_{r=1}^{s} q_r \leq m + s - 1,\\
& I_j \in \{0, 1\}, j \in E - \{o\},\\
& p_i \in \{0, 1\}, i = 1, \ldots, m,\\
& q_r \in \{0, 1\}, r = 1, \ldots, s,\\
& t_j \geq 0, j \in E,\\
& U \geq 0, V \geq 0,
\end{aligned}
$$

where E is the set of all extreme BCC-efficient DMUs, and quantities M, M_1 and M_2 are sufficiently large positive numbers.

The idea behind model 5 is as follows: the first three constraints guarantee that each feasible solution of model 9 is an optimal solution of model 5. The other constraints together with the objective function are designed in a way that the optimal solution of model 9 corresponds to the gradient of a WDH of T_v. Since we are minimizing $\sum_{i=1}^{m} p_i + \sum_{r=1}^{s} q_r + \sum_{j\in E-\{o\}} I_j$, and $I_j \in \{0,1\}$, $p_i \in \{0,1\}$ and $q_r \in \{0,1\}$, model 9 is directed toward finding optimal solutions with as many $I_j^* = 0$, $p_i^* = 0$ and $q_r^* = 0$ as possible. According to constraints (*), (**) and (***), this is equivalent to as many $t_j^* = 0$, $v_i^* = 0$, and $u_r^* = 0$ as possible or as many extreme directions as possible. The constraint $\sum_{i=1}^{m} p_i + \sum_{r=1}^{s} q_r \leq m+s-1$ guarantees that there certainly exist some zero components in the vector (U^*, V^*); i. e., the corresponding facet to the optimal solution of model 9 has certainly some directions of types (2) or (3).

Theorem 3.1 *Suppose that* DMU_o *is an extreme BBC-efficient DMU, and* (U^*, V^*, u_0^*) *is an optimal solution of model 9 in which* (U^*, V^*) *is not positive. If there exists at least one WDH passing through* DMU_o *with* $V \neq 0$*, then* $H_o^* : U^{*t}y - V^{*t}x + u_0^* = 0$ *is a WDH of* T_v*.*

See Appendix for proof.

Corollary 3.2 *Suppose that* DMU_o *is an extreme BCC-efficient DMU and the vector* (U^*, V^*, u_0^*) *is an optimal solution of model 9 that satisfies conditions of Theorem 3.1. Then* (U^*, V^*, u_0^*) *is an extreme optimal non-positive solution of model 5.*

We now propose a mixed binary program to determine the WDHs of T_v with $V = 0$; equivalently, for determining extreme optimal solutions of model 7. For this purpose, as in the above discussion, we must choose an optimal solution of model 7, say $\hat{U}$, such that $\hat{F} = \hat{H} \cap \hat{T}_v$ is a facet of $\hat{T}_v$, where $\hat{H} : \hat{U}^t y = 1$. Therefore, $\hat{U}$ should be chosen in such a way that the total directions of $\hat{F}$ be maximized; equivalently, we should choose an optimal solution of model 7 such that the number of zero elements in vector U and the number of members of $\hat{E}$ lying on $\hat{H}$ are totally maximized. To do so, we formulate the following model.

$$
\begin{aligned}
\min \quad & I_o = \sum_{r=1}^{s} q_r + \sum_{j\in \hat{E}-\{o\}} I_j && (10)\\
\text{s.t.} \quad & U^t y_o = 1, \\
& U^t y_j + t_j = 1, j \in \hat{E} - \{o\}, \\
& t_j - M I_j \leq 0, j \in \hat{E} - \{o\}, \\
& u_r - M_1 q_r \leq 0, r = 1, \ldots, s, \\
& I_j \in \{0,1\}, j \in \hat{E} - \{o\}, \\
& q_r \in \{0,1\}, r = 1, \ldots, s, \\
& t_j \geq 0, j \in \hat{E} - \{o\}, \\
& U \geq 0, V \geq 0.
\end{aligned}
$$

Theorem 3.3 *Suppose that* DMU_o *is an extreme BBC-efficient DMU, and* U^* *is an optimal solution of model 10. If there exists at least one WDH passing through* DMU_o *with* $V = 0$*, then* $H_o^* : U^{*t}y = 1$ *is a WDH of* T_v*.*

See Appendix for proof.

3.1 Algorithm

In this section, we use the structural characterization of the structure of FDWFs given in the previous section to propose an algorithm for finding all FDWFs of T_v. Our algorithm performs the following procedure for each extreme BCC-efficient DMU at each stage.

Consider the extreme BCC-efficient observed unit, DMU$_o$. We perform the following main step for DMU$_o$. Note that the main step comprises two phases. First, phase 1 is implemented for DMU$_o$ to determine all existing WDHs of T_v with $V \neq 0$ passing through DMU$_o$ and their corresponding FDWFs. In Phase 2 we determine all existing WDHs of T_v with $V = 0$ and their pertaining FDWFs containing DMU$_o$.

Main step.

Phase 1: Evaluate DMU_o by model 9. If there exists at least one WDH with $V \neq 0$ containing DMU_o, then the optimal solution of model 9 is the gradient of a WDH with $V \neq 0$ passing through DMU_o. If the optimal objective value of model 9 is strictly greater than $|E|$, i.e., the hyperplane corresponding to the optimal solution has strictly less than $m+s-1$ extreme directions, then there is no more WDH with $V \neq 0$ passing through DMU_o. Therefore, Phase 1 iterates the following steps until the optimal objective value of model 9 is strictly greater than $|E|$.

Suppose that I_o^* and (U^*, V^*, u_0^*) are the optimal objective and optimal solution of model 9, respectively. Let $H_o^* : U^{*t}y - V^{*t}x + u_0^* = 0$, and $J_o^* = \{j : I_j^* = 0\}$, $Q_o^* = \{r : q_r^* = 0\}$ and $P_o^* = \{i : p_i^* = 0\}$. The set J_o^* contains the indices of all other extreme BCC-efficient DMUs besides DMU_o lying on H_o^*, the set Q_o^* contains the indices of vector U^* that are zero, and P_o^* is the set of indices i associated to zero elements of vector V^*. We take H_o^* and

$$F_o^* = \Big\{(x,y) \mid (x,y) = \sum_{j\in J_o^*} \lambda_j (x_j, y_j) + \sum_{i\in P_o^*} \mu_i (e_i, 0) + \sum_{r\in Q_o^*} \beta_r (0, -e_r),$$
$$\sum_{j\in J_o^*} \lambda_j = 1, \lambda_j \geq 0, j \in J_o^*, \mu_i \geq 0, i \in P_o^*, 0 \leq \beta_r \leq \delta_r, \text{for some } \delta_r, r \in Q_o^*\Big\} = H_o^* \cap T_v$$

as a WDH with $V \neq 0$ and FDWF of T_v, respectively. Next, we add

$$\sum_{j\in J_o^*} I_j + \sum_{i\in P_o^*} p_i + \sum_{r\in Q_o^*} q_r \geq 1, \tag{11}$$

to the constraints of model 9 to prevent obtaining a repetitive hyperplane and evaluate DMU_o by model 9 again. If there exists another WDH with $V \neq 0$ besides H_o^* passing through DMU_o, the model 9 with constraint 11 gives the gradient of another WDH with $V \neq 0$ passing through DMU_o as an alternative optimal solution. We keep this WDH with $V \neq 0$ and construct the new set J_o^*, P_o^* and Q_o^* corresponding to the new WDH. If there is no more WDH with $V \neq 0$ except H_o^* passing through DMU_o, Phase 1 is terminated for DMU_o.

Suppose that steps of Phase 1 are repeated t times for DMU_o. Then $t-1$ WDHs with $V \neq 0$ are determined. Note that in the final step, model 9 will have exactly $t-1$ new added constraints corresponding to $t-1$ WDHs. Once Phase 1 is terminated for DMU_o, all WDHs of T_v with $V \neq 0$ passing through DMU_o and all other extreme BCC-efficient DMUs besides DMU_o lying on these hyperplanes are identified. This set is denoted by E_o^1.

Theorem 3.4 below shows that Phase 1 produces all WDHs with $V \neq 0$ of T_v.

Theorem 3.4 *All WDHs with $V \neq 0$ passing through DMU_o are determined before Phase 1 is terminated for DMU_o.*

See Appendix for proof.

If $o \notin \hat{E}$, then all WDHs for DMU_o are determined in Phase 1. Otherwise, there exists at least one WDH with $V = 0$ of T_v passing through DMU_o. To find these WDHs, we implement Phase 2.

Phase 2: If $s = 1$, i.e., we have only one output variable, then the WDH is $y = \max_{j\in\hat{E}} y_j$. Otherwise, evaluate DMU_o by model 10. If there exists at least one WDH with $V = 0$ containing DMU_o, then the optimal solution of model 10 is the gradient of a WDH with $V = 0$ passing through DMU_o. If the optimal objective value of model 10 is strictly greater than $|E|$, i.e., the hyperplane corresponding to the optimal solution has strictly less than $m+s-1$ extreme directions, there is no more WDH with $V = 0$ passing through DMU_o. The following steps are therefore iterated in Phase 2 until the optimal value of model 10 is strictly greater than $|E|$.

Suppose that $\bar{I}_o$ and $\bar{U}$ are the optimal objective and optimal solution of model 10, respectively. Let $\bar{H}_o : \bar{U}^t y = 1$, $\bar{J}_o = \{j : \bar{I}_j = 0\}$ and $\bar{Q}_o = \{r : \bar{q}_r = 0\}$. The set $\bar{J}_o$ contains the indices of all other extreme

BCC-efficient DMUs besides DMU$_o$ lying on $\bar{H}_o$, and $\bar{Q}_o$ is the set of indices r associated to zero elements of vector $\bar{U}$. We take $\bar{H}_o$ and

$$\bar{F}_o = \Big\{(x,y) \mid (x,y) = \sum_{j\in\bar{J}_o} \lambda_j(x_j,y_j) + \sum_{i=1}^{m} \mu_i(e_i,0) + \sum_{r\in\bar{Q}_o} \beta_r(0,-e_r),$$
$$\sum_{j\in\bar{J}_o} \lambda_j = 1, \lambda_j \geq 0, j\in\bar{J}_o, \mu_i \geq 0, i \in \bar{P}_o, 0 \leq \beta_r \leq \delta_r, \text{for some } \delta_r, r\in\bar{Q}_o\Big\} = \bar{H}_o \cap T_v,$$

as a WDH with $V=0$ and FDWF of T_v, respectively. Next, we add the following constraint to the constraints of model 10:

$$\sum_{j\in\bar{J}_o} I_j + \sum_{r\in\bar{Q}_o} q_r \geq 1. \tag{12}$$

Suppose that steps of Phase 2 are repeated k times for DMU$_o$. Then $k-1$ WDHs with $V=0$ are determined. Note that in the final step, model 10 will have exactly $k-1$ new added constraints corresponding to $k-1$ WDHs. Once Phase 2 is terminated for DMU$_o$, all WDHs of T_v, with $V=0$ passing through DMU$_o$ and all other extreme BCC-efficient DMUs besides DMU$_o$ lying on these hyperplanes are identified. We denote this set by E_o^2. Define $E_o = E_o^1 \cup E_o^2$. Then E_o is the set of all extreme BCC-efficient DMUs besides DMU$_o$ that are at least on one FDWF with DMU$_o$.

Similarly to Theorem 3.4, we can prove that Phase 2 determines all WDHs with $V=0$ of T_v.

Theorem 3.5 *All WDHs with $V=0$ passing through DMU$_o$ are determined before Phase 2 is terminated for DMU$_o$.*

See Appendix for proof.

The main step must be performed for all extreme BCC-efficient DMUs. To avoid obtaining the gradients of iterated WDHs determined in the implementation of the main step for DMU$_o$, the constraint $I_o = 1$ must always be added to the constraints of model 9 and 10 in all subsequent stages where we implement the main step for other extreme BCC-efficient DMUs. Having implemented the main step for all extreme BCC-efficient DMUs, Theorems 3.4 and 3.5 guarantee that all WDHs of T_v will be determined.

Remark 2 *Note that since the dimension of each FDWF of T_v is $m+s-1$, Phases 1 and 2 of the main step for DMU$_o$ is repeated at most $\binom{|E_o^1|+m+s-r}{m+s-1}$ and $\binom{|E_o^2|+s-r}{s-1}$, respectively; where r is the number of DMUs that have been evaluated before DMU$_o$. In most cases, however, Phase 1 and Phase 2 are terminated in fewer steps. For example, as shown in the next section, the algorithm is repeated only 2 times in Phase 1 for unit D while the upper bound is $\binom{5}{2} = 10$.*

3.2 Summary of the algorithm

Suppose that we have n observed DMUs, DMU$_j$, $j = 1,\ldots,n$. Let
$E = \{j : DMU_j$ is extreme BCC-efficient$\}$ and $\hat{E} = \{j \mid y_j$ is extreme BCC-efficient of $\hat{T}_v\}$.

Set $E_T = \emptyset$ and $S = \emptyset$.

Step 1 Choose $DMU_p \in E \setminus E_T$ and set $E_p = \emptyset$ and $S_p = \emptyset$.

Step 2 Evaluate DMU_p using model 9. If the optimal objective value of model 9 is strictly greater than $|E|$, then go to Step 5. Otherwise go to Step 3.

Step 3 For the optimal solution (U^*, V^*, u_0^*) of model 9, set $J_p^* = \{j \in E : I_j^* = 0\}$, $P_p^* = \{i : p_i^* = 0\}$ and $Q_p^* = \{r : q_r^* = 0\}$. Put the hyperplane $H_p^* : U^{*t}y - V^{*t}x + u_0^* = 0$ into set S_p. Set $E_p = E_p \cup J_p^*$, $S_p = S_p \cup \{H_p^*\}$ and $S = S \cup \{S_p\}$.

Step 4 Add the constraint

$$\sum_{j\in J_p^*} I_j + \sum_{i\in P_p^*} p_i + \sum_{r\in Q_p^*} q_r \geq 1,$$

to the constraints of model 9 and go to Step 2.

Step 5 If $p \in \hat{E}$, go to Step 6; otherwise $E_T = E_T \setminus \{DMU_p\}$ and go to Step 9.

Step 6 Evaluate DMU_p using model 10. If the optimal value of model 10 is strictly greater than $|E|$, put $E_T = E_T \setminus \{DMU_p\}$ and go to Step 9. Otherwise go to Step 7.

Step 7 For the optimal solution $\bar{U}$ of model 10, set $\bar{J}_p = \{j : \bar{I}_j = 0\})$ and $\bar{Q}_p = \{r : \bar{q}_r = 0\}$. Put the hyperplane $\bar{H}_p : \bar{U}^t y = 1$ into set S_p. Set $E_p = E_p \cup \bar{J}_p$, $S_p = S_p \cup \{\bar{H}_p\}$ and $S = S \cup \{S_p\}$.

Step 8 Add the constraint

$$\sum_{j \in \bar{J}_p} I_j + \sum_{r \in \bar{Q}_p} q_r \geq 1,$$

to the constraints of model 10 and go to Step 6.

Step 9 If $E = E_T$, the algorithm is terminated; otherwise, add constraint $I_p = 1$ to the constraints of models 9 and 10, and go to step 1.

The flowchart of this algorithm is depicted in Figure 3.

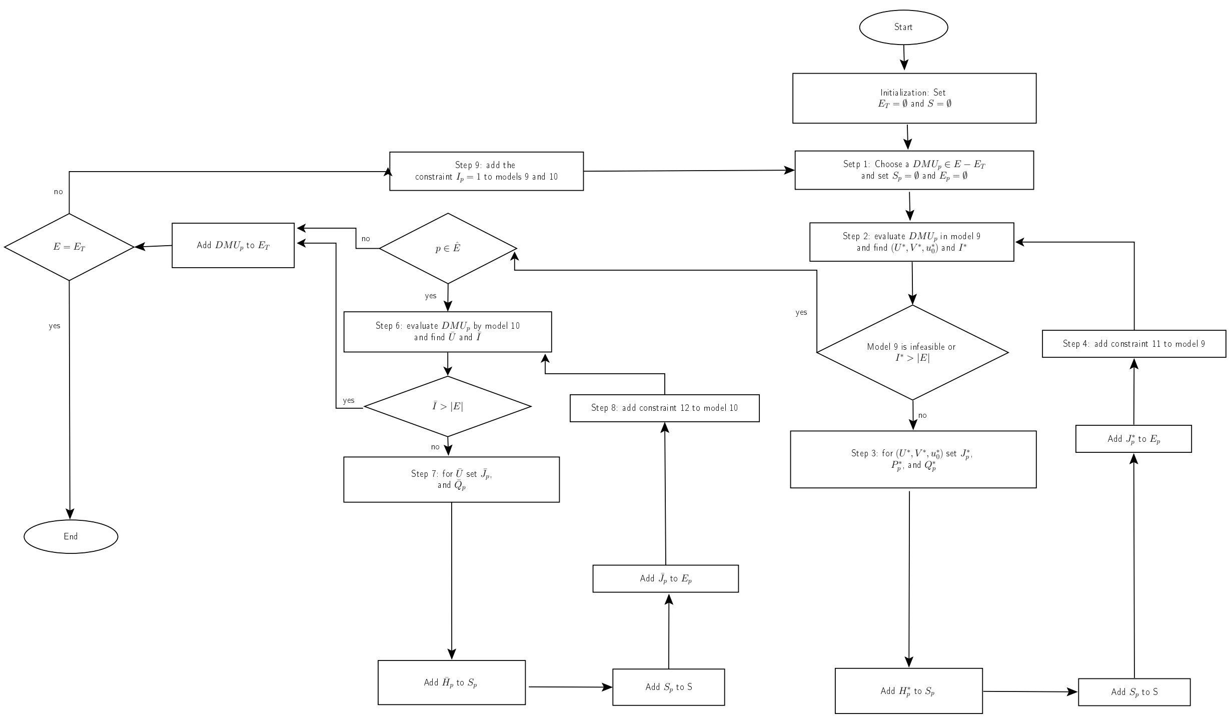

Figure 3: Flowchart of the proposed algorithm.

Remark 3 *To implement our algorithm in Phase 1, we use the input oriented BCC form to determine WDHs with $V \neq 0$. In Phase 2 we take the output oriented of BCC model to determine WDHs with $V = 0$. We call this approach $I - O$ method. The algorithm can be implemented in the reverse order, by replacing V by U, i.e., starting with the output oriented of BCC model to find WDHs with $U \neq 0$ and then in Phase 2 obtain those with $U = 0$ using the input oriented of BCC model. This latter approach we call $O - I$ method.*

Remark 4 *The above remark gives rise to the following question: which method should we take for a given set of data? The rule of thumb is to choose the approach that leads to solving the lesser number of mixed binary programs with a larger number of binary variables. Therefore, if the number of input variables is less than the number of output variables, the $I - O$ method is generally more preferable. Otherwise, we suggest the $O - I$ method.*

4 Illustrative example

In this section, we implement the $I-O$ approach for the data of four hypothetical DMUs listed in Table 1. Units A, B, C, and D use two inputs to produce one output. The PPS, T_v, constructed by those DMUs are depicted in Figure 4. As mentioned earlier, to find all WDHs, it suffices to perform our algorithm for each of the extreme BCC efficient DMUs of T_v in Phase 1 and each of the extreme BCC efficient DMUs of $\hat{T}_v$ in Phase 2. The WDHs and associated sets E_j obtained, using our algorithm, for each of extreme BCC efficient DMUs of T_v are documented in Table 2. The details of our algorithm for unit D are given below. Details for other units can be similarly obtained. It can be easily tested that $E = \{A, B, C, D\}$ and $\hat{E} = \{C, D\}$. Therefore, Phase 2 should be implemented only for units C and D.

Table 1: Data of inputs and output of DMUs.

DMU	A	B	C	D
x_1	3	10	15	6
x_2	10	3	6	15
y	6	6	9.5	9.5

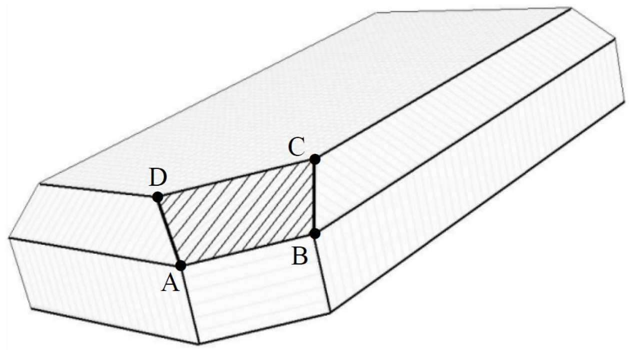

Figure 4: The PPS constructed with units listed in Table 1.

Let $E_T = \emptyset$ and $S = \emptyset$.

Step 1. Put $D \in E - E_T$ and evaluate the unit using model 9. $I_D^{*1} = |E| = 4$, and $(v_1^*, v_2^*, u_1^*, u_0^*) = (0.1667, 0, 0.1429, -0.3571)$ is the optimal solution of model 9. Since (v_1^*, v_2^*, u_1^*) is not positive, $H_1 : 0.1429y_1 - 0.1667x_1 - 0.3571 = 0$ is the WDH of T_v and we have $S = \{H_1\}$. Furthermore, $t_A^* = v_2^* = 0$ thereby $J_D^* = \{A\}$, $P_D^* = \{I_2\}$, $Q_D^* = \emptyset$, $J_D^{*c} = \{B, C\}$, $P_D^{*c} = \{I_1\}$, and $Q_D^{*c} = \{O\}$. Set the following constraint

$$I_A + P_2 \geq 1.$$

Step 2. Add the above constraint to the constraints of model 9 and evaluate unit D by model 9 again. We have $I_D^{*2} = 5 > |E|$. Thus, Phase 1 is terminated for unit D.

Step 3. Since $D \in \hat{E}$, we must implement Phase 2 for unit D. Given that there is only one output variable, the WDHs with $V = 0$ of T_v is $H_2 : y = \max_{j \in \hat{E} = \{C,D\}} \{y_j\} = y_C = y_D = 9.5$.

Add constraint $I_D = 1$ to the constraints of model 9 and move to the next stage, i.e., implement algorithm for the other BCC efficient DMUs. This can prevent the algorithm from obtaining the gradients of iterated WDHs, i. e., H_1, determined in the implementation of the algorithm for unit D.

We can summarize the results obtained through the implementation of our algorithm for other extreme BCC efficient DMUs as follows:

- The algorithm obtains only $H_3 : 0.1429y - 0.1667x_2 - 0.3571 = 0$ before its termination for unit C. We then add constraint $I_C = 1$ and go to the next stage.

Table 2: The summary of the results.

DMU	The WDHs and its corresponding FDWFs passing through the unit	E_j
A	$H_1: 0.1429y_1 - 0.1667x_1 - 0.3571 = 0$ $H_4: 0.7692x_1 + 0.7692x_2 - 1 = 0$ $H_6: 0.3333x_1 + 1 = 0$ $J_A^{*6} = \emptyset$ $F_1 = \{(x_1, x_2, y) \mid (x_1, x_2, y) = \lambda_1 A + \lambda_2 D + \mu(0,1,0), \lambda_1 + \lambda_2 = 1, \lambda_j \geq 0, j = 1,2, \mu \geq 0\} =$ $\{(x_1, x_2, y) \mid (x_1, x_2, y) = (3\lambda_1 + 6\lambda_2, 10\lambda_1 + 15\lambda_2 + \mu, 6\lambda_1 + 9.5\lambda_2), \lambda_1 + \lambda_2 = 1, \lambda_j \geq 0, j = 1,2, \mu \geq 0\} = H_1 \cap T_v.$ $F_4 = \{(x_1, x_2, y) \mid (x_1, x_2, y) = (3\lambda_1 + 10\lambda_2, 10\lambda_1 + 13\lambda_2, 6 - \beta)\lambda_1 + \lambda_2 = 1, \lambda_j \geq 0, j = 1,2, 0 < \beta < \delta_1, \text{for some } \delta_1\} = H_4 \cap T_v.$ $F_6 = \{(x_1, x_2, y) \mid (x_1, x_2, y) = (3, 10 + \mu, 6 - \beta), \mu \geq 0, 0 < \beta < \delta_2, \text{for some } \delta_2\} = H_6 \cap T_v.$	$J_A^{*1} = \{D\}$ $J_A^{*4} = \{B\}$ $E_A = \{B, D\}$
B	H_4, $H_5: -0.3333x_2 + 1 = 0$, and $H_3: 0.1429y - 0.1667x_2 - 0.3571 = 0$ F_4, $F_5 = \{(x_1, x_2, y) \mid (x_1, x_2, y) = (10 + \mu, 3, 6 - \beta), \mu \geq 0, 0 < \beta < \delta_3, \text{for some } \delta_3\} = H_5 \cap T_v$, and $F_3 = \{(x_1, x_2, y) \mid (x_1, x_2, y) = (10\lambda_1 + 15\lambda_2 + \mu, 3\lambda_1 + 6\lambda_2, 6\lambda_1 + 9.5\lambda_2), \lambda_1 + \lambda_2 = 1, \lambda_j \geq 0, j = 1,2, \mu \geq 0\} = H_3 \cap T_v.$	$J_B^{*5} = \emptyset$ $J_B^{*3} = \{C\}$ $E_B = \{A, C\}$
C	H_3 and $H_2: y = 9.5$. F_3 and $F_2 = \{(x_1, x_2, y) \mid (x_1, x_2, y) = (15\lambda_1 + 6\lambda_2 + \mu_1, 6\lambda_1 + 15\lambda_2 + \mu_2, 9.5), \lambda_1 + \lambda_2 = 1, \lambda_j \geq 0, j = 1,2, \mu_1, \mu_2 \geq 0\} = H_2 \cap T_v.$	$E_C = \{B, D\}$
D	H_1, H_2 F_1, F_2	$E_D = \{C, A\}$

- The WDHs $H_4 : 0.7692x_1 + 0.7692x_2 - 1 = 0$ and $H_5 : -0.3333x_2 + 1 = 0$ are obtained from the implementation of our algorithm for unit B. The WDH H_3 also passes through unit B. However, adding the constraint $I_C = 1$ to the constraints of model 9 prevents the algorithm from generating this WDH when we run the procedure for D. We add constraint $I_B = 1$ to the constraints of model 9 and then run our method for unit A.
- The WDH $H_6 : 0.3333x_1 + 1 = 0$ is obtained. The WDHs H_4 and H_1 also pass through unit A. However, adding the constraints $I_B = 1$ and $I_C = 1$ in the pervious stages to the constraints of model 9 prevents the algorithm from generating these iterated WDHs.

Appendix

Proof of Theorem 2.1. Since (U^*, V^*, u_0^*) is an extreme optimal solution of model 5, there exist $m+s$ linearly independent constraints of model 5 binding at (U^*, V^*, u_0^*). Suppose, without loss of generality, that $U^t y_j - V^t x_j + u_0 \leq 0, j = 2, \ldots, k-1$, $v_i \geq 0, i = 1, \ldots, t$, and $u_r \geq 0, r = 1, \ldots, l$ are these $m+s$ linearly independent constraints. Then the following matrix is row full rank:

$$\begin{pmatrix} -x_o & y_o & 1 \\ -x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ -x_k & y_k & 1 \\ e_1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ e_t & 0 & 0 \\ 0 & e_1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & e_l & 0 \end{pmatrix}$$

and it is row equivalent to the following matrix:

$$\begin{pmatrix} -x_o & y_o & 1 \\ x_o - x_2 & y_2 - y_o & 0 \\ \vdots & \vdots & \vdots \\ x_o - x_k & y_k - y_o & 0 \\ -e_1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ -e_t & 0 & 0 \\ 0 & -\kappa e_1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & -\kappa e_l & 0 \end{pmatrix}$$

Thus the set $D = \{DMU_j - DMU_o\}_{j=2}^{k} \bigcup \left\{\widehat{DMU}_i - DMU_o\right\}_{i=1}^{t} \bigcup \left\{\widetilde{DMU}_p - DMU_o\right\}_{p=1}^{l}$ is linearly independent where $\widehat{DMU}_i = (x_o + e_i, y_o), i = 1, \ldots, t$ and $\widetilde{DMU}_p = (x_o, y_o - \kappa e_p), p = 1, \ldots, l$. Obviously, $\widehat{DMU}_i, i = 1, \ldots, t$ and $\widetilde{DMU}_p, p = 1, \ldots, l$ lie on H^*. Since $y_{ro} > 0, r = 1, \ldots, s$, there exists a $0 \leq \kappa \leq \min_{r \in \{1,\ldots,s\}} y_{ro}$. Having chosen such κ and using disposability axiom $\widehat{DMU}_i$ and $\widetilde{DMU}_p \in T_v$, for $i = 1, \ldots, t, p = 1, \ldots, l$ and hence they all lie on $F^* = H^* \cap T_v$. Thus, there exist $m+s$ affinely independent DMUs (observed or virtual) lying on F^*. Using Definition 3, F^* is an FDWF of T_v.

We can similarly prove that the hyperplanes corresponding to extreme optimal and non-positive solutions of model 6 are WDHs passing through DMU$_o$. $\square$

Proof of Theorem 2.2. Suppose that $H_o^*: U^{*t}y - V^{*t}x + u_0^* = 0$ is a WDH of the PPS. Because $V^{*t}x_o \neq 0$ (res. $U^{*t}y_o \neq 0$), we assume, without loss of generality, that H_o^* is normalized, i. e., $V^{*t}x_o = 1$ (res. $U^{*t}y_o = 1$). Therefore

$$U^{*t}y_o + u_0^* = 1(res. V^{*t}x_o - u_0^* = 1). \tag{13}$$

Using the definition of WDH, H_o^* is a weak and supporting hyperplane. Thus

$$U^{*t}y_j - V^{*t}x_j + u_0^* \leq 0, j = 1, \ldots, n \tag{14}$$

and

$$(U^*, V^*) \geq 0. \tag{15}$$

Using (13), (14), and (15), (U^*, V^*, u_0^*) is an optimal solution of model 5 (res. model 6). It then suffices to show that (U^*, V^*, u_0^*) is also extreme. Suppose that (U^*, V^*, u_0^*) is not extreme. Then there exists an extreme optimal solution of model 5 (model 6), say $(\bar{U}, \bar{V}, \bar{u}_0)$, such that all constraints that are binding at (U^*, V^*, u_0^*) are also binding at $(\bar{U}, \bar{V}, \bar{u}_0)$. Thus, each DMU lying on H_o^* also lies on the hyperplane corresponding to $(\bar{U}, \bar{V}, \bar{u}_0)$. Since H_o^* is WDH of T_v, there exist $m + s$ affinely independent DMUs (observed and virtual) lying on H_o^*. Therefore those $m + s$ affinity independent DMUs also lie on the hyperplane corresponding to $(\bar{U}, \bar{V}, \bar{u}_0)$. Thus, there exist two different hyperplanes passing through one affinely independent set with $m + s$ elements. This is a contradiction. □

Proof of Theorem 2.3.

Part 1. Since $\bar{H} \cap T_v$ is a convex set, the point $(x_o, y_o) + \lambda d_k = (1 - \lambda)(x_o, y_o) + \lambda(x_k, y_k)$ lies on $\bar{H} \cap T_v$ for each λ, $0 \leq \lambda \leq 1$. Therefore d_k is a feasible direction of $\bar{H} \cap T_v$ at point DMU$_o$. Suppose that d_k is not extreme. Then there exist two distinct feasible directions of $\bar{H} \cap T_v$ at DMU_o namely, d_α and d_β, such that $d_k = \eta d_\alpha + (1 - \eta)d_\beta$, $\eta > 0$. Since d_α and d_β are two feasible directions of $\bar{H} \cap T_v$ at point DMU_o, so are $d_\alpha = (x_\alpha - x_o, y_\alpha - y_o)$ and $d_\beta = (x_\beta - x_o, y_\beta - y_o)$ for two points (x_α, y_α) and (x_β, y_β) in the set $\bar{H} \cap T_v$. Therefore, $(x_o, y_o) = \eta(x_\alpha, y_\alpha) + (1 - \eta)(x_\beta, y_\beta)$, for some $\eta > 0$. On the other hand, (x_o, y_o) is an extreme point of the set $\bar{H} \cap T_v$. This is a contradiction.

Part 2. Using the fact that $d_k = (e_k, 0)$ and $(\bar{U}, \bar{V})$ with $\bar{v}_k = 0$ are orthogonal, and the disposability axiom, $(\tilde{x}, \tilde{y}) + \lambda(e_k, 0) \in \bar{H} \cap T_v$ for any point $(\tilde{x}, \tilde{y})$ in the set $\bar{H} \cap T_v$ and any positive scalar λ. Therefore, $d_k = (e_k, 0)$ is a recession direction of the set $\bar{H} \cap T_v$. It then suffices to show that it is also extreme. Suppose this is not the case. Then, there exist two distinct recession directions of $\bar{H} \cap T_v$ namely, $\bar{d}$ and $\tilde{d}$, such that $(e_k, 0) = \bar{\lambda}\bar{d} + \tilde{\lambda}\tilde{d}$, $\bar{\lambda}$, $\tilde{\lambda} > 0$. Let $\bar{d} = (\bar{d}_x, \bar{d}_y)$ and $\tilde{d} = (\tilde{d}_x, \tilde{d}_y)$ where $\bar{d}_x, \tilde{d}_x \in R^{m \geq 0}$ and $\bar{d}_y, \tilde{d}_y \in R^{s \geq 0}$. Thus, we have $\bar{d}_y, \tilde{d}_y = 0$, $\bar{d}_x = \bar{\alpha} e_k$, $\tilde{d}_x = \tilde{\alpha} e_k$ and $\bar{\alpha}, \tilde{\alpha} > 0$. This is a contradiction.

Part 3. Using the fact that $d_k = (0, -e_k)$ and $(\bar{U}, \bar{V})$ with $\bar{u}_q = 0$ are orthogonal, and the disposability axiom, $(x_o, y_o) + \lambda d_k \in \bar{H} \cap T_v$ for any $\lambda, 0 \leq \lambda \leq y_{ko}$. If $y_{ko} > 0$, then d_k is a feasible direction of $\bar{H} \cap T_v$ at point DMU_o. Otherwise $y_{ko} = 0$ and d_k is not a feasible direction at DMU$_o$. Similarly to the proof of part 2, we can prove that $d_k = (0, -e_k)$ is also extreme. □

Proof of Theorem 3.1. Suppose that (U^*, V^*, u_0^*) is an optimal solution of model 9. Consider the following model

$$\begin{aligned} \max \quad & U^t y_o + u_0 \\ \text{s.t.} \quad & V^t x_o = 1, \\ & U^t y_j - V^t x_j + u_0 \leq 0, j \in E, \\ & U \geq 0, V \geq 0. \end{aligned} \tag{16}$$

Model 16 is model 5, with constraints being restricted to $j \in E$. Obviously, (U^*, V^*, u_0^*) is an optimal solution of model 16. On the other hand, the optimal solutions of models 5 and 16 are equal. Therefore, using Theorem 2.1, it suffices to prove that the optimal solution of model 9 is an extreme (basic) optimal solution of model 16 in evaluating DMU_o. Suppose that (U^*, V^*, u_0^*) is not an extreme optimal solution of model 16.

Let $S = \{\left(\hat{U}^1, \hat{V}^1\right), \dots, \left(\hat{U}^l, \hat{V}^l\right)\}$ be the set of all gradients of WDHs passing through DMU_o. Using Theorem 2.2, the vectors $\left(\hat{U}^r, \hat{V}^r, \hat{u_0^r}\right), r = 1, \dots, l$ are all the extreme optimal solutions of model 16 with some zero elements. Since (U^*, V^*, u_0^*) is not an extreme optimal solution of model 16, (U^*, V^*) can be represented as a convex combination of members of S. In other words, there exist $\hat{\alpha}^{i_1}, \dots, \hat{\alpha}^{i_p}$ such that

$$(U^*, V^*) = \sum_{j=1}^{p} \hat{\alpha}^{i_j} \left(\hat{U}^{i_j}, \hat{V}^{i_j}\right), \tag{17}$$

$$\sum_{j=1}^{p} \hat{\alpha}^{i_j} = 1, 0 < \hat{\alpha}^{i_j} < 1, j = 1, \dots, p,$$

where $\hat{S} = \{\left(\hat{U}^{i_1}, \hat{V}^{i_1}\right), \dots, \left(\hat{U}^{i_p}, \hat{V}^{i_p}\right)\} \subseteq S$. Since (U^*, V^*, u_0^*) is not extreme $\mathrm{card}\{\hat{\mathrm{S}}\} \geq 2$. Next, we prove that for any $k \in \{1, \dots, p\}$ the sets of constraints binding at (U^*, V^*, u_0^*) and $\left(\hat{U}^{i_k}, \hat{V}^{i_k}, \hat{u}_0^{i_k}\right)$ are equal. Fix $k \in \{1, \dots, p\}$. Let A and B be respectively the sets of all constraints binding at (U^*, V^*, u_0^*) and $\left(\hat{U}^{i_k}, \hat{V}^{i_k}, \hat{u}_0^{i_k}\right)$. It is clear that $A \subseteq B$. Suppose that $B \nsubseteq A$. Then, in addition to constraints binding at (U^*, V^*, u_0^*), there exists at least one more constrict binding at $\left(\hat{U}^{i_k}, \hat{V}^{i_k}, \hat{u}_0^{i_k}\right)$ but not at (U^*, V^*, u_0^*). Note that an extra binding condition $(t_j = 0)$ implies that an extra $I_j = 0$. Thus $\left(\hat{U}^{i_k}, \hat{V}^{i_k}, \hat{u}_0^{i_k}\right)$ is a feasible solution of model 9 with an objective value less than I_o^*. This is a contradiction. Therefore, $A = B$. Thus, there exist some $m + s$ linearly independent constraints of model 16 that are commonly binding at each extreme optimal solution $\left(\hat{U}^{i_j}, \hat{V}^{i_j}, \hat{u}_0^{i_j}\right)$, for $j = 1, \dots, p$. This is a contradiction. □

Proof of Theorem 3.3. By rewriting model 9 based on model 6, we obtain the output oriented version of model 9 in which the first two constraints of model 9 are replaced by $U^t y_o = 1$ and $V^t x_o - u_0 = 1$, respectively. Similarly to Theorem 3.1, we can prove that if (U^*, V^*, u_0^*) is an optimal solution of the output version of model 9 in evaluating DMU_o and also there exists at least one WDH passing through DMU_o with $U \neq 0$, then $H_o^* : U^{*t} y - V^{*t} x + u_0^* = 0$ is a WDH of T_v. Since U^* is an optimal solution of model 10, $(U, V, u_0) = (U^*, 0, -1)$ is an optimal solution of the output oriented version of model 9. On the other hand, since there exists at least one WDH passing through DMU_o with $V = 0$, there exists at least one WDH passing through DMU_o with $U \neq 0$. Thus, $H_o^* : U^{*t} y = 1$ is a WDH of T_v. □

Proof of Theorem 3.4. Suppose that Phase 1 is terminated for DMU_o, but there still exists some WDHs with $V \neq 0$ of T_v passing through DMU_o. Suppose that $\bar{H}_o : \bar{U}^t y - \bar{V}^t x + \bar{u}_0 = 0$ in which $\bar{V} \neq 0$ is the remaining WDH of T_v passing through DMU_o. Obviously, $(\bar{U}, \bar{V}, \bar{u}_0)$ is a feasible solution of model 9. Suppose that the objective value at $(\bar{U}, \bar{V}, \bar{u}_0)$ is $\bar{I}_o = |E| + m + s - 1 - k$. Since $\bar{H}_o$ is a WDH, we have $k \geq m + s - 1$ and hence $\bar{I}_o \leq |E|$. Thus, model 9 is feasible and its objective value is less than or equal to $|E|$. This is a contradiction. □

Proof of Theorem 3.5. Suppose that Phase 2 is terminated for DMU_o, but there still exists some WDHs with $V = 0$ of T_v passing through DMU_o. Suppose that $\bar{H}_o : \bar{U}^t y = 1$ is the remaining WDH of T_v passing through DMU_o. Obviously, $\bar{U}$ is a feasible solution of model 10. Suppose that the objective value at $\bar{U}$ is $\bar{I}_o = |E| + s - 1 - k$. Since $\bar{H}_o$ is a WDH, we have $k \geq s - 1$ and hence $\bar{I}_o \leq |E|$. Thus, the objective value of model 10 is less than or equal to $|E|$. This is a contradiction. □

References

[1] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. European Journal of Operational Research. 1978; 2:429–444.

[2] Banker RD, Charens A, Cooper WW. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science. 1984; 30:1078–1092.

[3] Yu G, Wei Q, Brockett P, Zhou L. Construction of all DEA efficient surfaces of the production possibility set under the generalized data envelopment analysis mode. European Journal of Operational Research. 1996; 95:491–510.

[4] Olesen O, Petersen N. Indicators of ill-conditioned data sets and model misspecification in data envelopment analysis. Management Science. 1996; 42:205–219.

[5] Olesen O, Petersen N. Identification and use of efficient faces and facets in DEA. Journal of Productivity Analysis. 2003; 20:323–360.

[6] Jahanshahloo GR, Lotfi FH, Rezai HZ, Balf FR. Finding strong defining hyperplanes of production possibility set. European Journal of Operational Research. 2007; 177:42–54.

[7] Jahanshahloo GR, Shirzadi A, Mirdehghan M. Finding strong defining hyperplanes of PPS using multiplier form. European Journal of Operational Research. 2009; 194:933–938.

[8] Davtalab-Olyaie M, Roshdi I, Jahanshahloo G, Asgharian M. Characterizing and finding full dimensional efficient facets in dea: A variable returns to scale specification. journal of the Operational Research Society (To appear). 2013.

[9] Aparicio J, Ruiz JL, Sirvent I. Finding closest targets and minimum distance to the Pareto-efficient frontier in DEA technologies. Journal of Productivity Analysis. 2007; 28:209–218.

[10] Chang KP, Guh YY. Linear production functions and the data envelopment analysis. European Journal of Operational Research. 1991; 52:215–223.

[11] Jahanshahloo GR, Lotfi FH, Shoja N, Sanei M, Tohidi G. Sensitivity and stability analysis in data envelopment analysis. Journal of Operational Research Society. 2005; 56:342–345.

[12] Boljuncic V. Sensitivity analysis of an efficient DMU in DEA model with variable returns to scale (VRS). Journal of Productivity Analysis. 2006; 25:173–192.

[13] Wei Q, Yan H, Hao G. Characteristics and structures of weak efficient surfaces of production possibility. Mathematical Analysis and Applications. 2007; 327:1055–1074.

[14] Jahanshahloo GR, Lotfi FH, Akbarian D. Finding weak defining hyperplanes of PPS of the BCC model. Applied Mathematical Modelling. 2010; 34:3321–3332.

[15] Murty K. Linear programming. New York: Wiley; 1983.

[16] Bazaraa MS, Jarvis JJ, Sherali HD. Linear programming and network flows. New Jersey: John Wiley & Sons; 2005.

[17] Cooper WW, Seiford LM, Tone K. Data Envelopment Analysis: A comprehensive text with models, applications, references and DEA-solver software (2th ed.). Berlin: Springer; 2007.

[18] Charnes A, Cooper WW, Thrall RM. Classifying and characterizing efficiencies and inefficiencies in data envelopment analysis. Operations Research Letters. 1986; 5:105–110.