Available online at http://ijim.srbiau.ac.ir/
Int. J. Industrial Mathematics (ISSN 2008-5621)
Vol. 10, No. 4, 2018 Article ID IJIM-01034, 11 pages
Research Article

Science and Research Branch (IAU)

Weak Disposability in Integer-Valued Data Envelopment Analysis

A. Amirteimoori *†, M. Maghbouli ‡

Received Date: 2017-02-26 Revised Date: 2017-10-22 Accepted Date: 2018-02-01

Abstract

Conventional data envelopment analysis (DEA) models normally assume all inputs and outputs are real valued and continuous. However in most application- related problems some inputs and outputs can only take integer values, also, both desirable and undesirable outputs can be generated (e.g., the number of traffic accidents and deaths in a transportation system). In this paper the effect of undesirable outputs in integer DEA model is discussed. The proposed model distinguishes weak disposability of outputs imposing non-uniform abatement factor. Compared with radial models, a non radial model that directly deals with slacks is developed to calculate efficiency for integer-valued data set. An empirical application is used to illustrate the approach.

Keywords : Data Envelopment Analysis (DEA); Weak-Disposability; Undesirable factors; Integer-valued data; DMU; Efficiency.

1 Introduction

For a long time, Data Envelopment Analysis (DEA) has been serving as a methodology to evaluate the performance of various decision making units (DMU) that consumes multiple inputs to generate multiple outputs. Conventional DEA models have been created according to application, radial models as CCR (Charnes et.al [3]) and BCC (Banker et.al [1]) or non-radial models as SBM (Tone, [20]) and so on. These models take real-valued desirable inputs and outputs. A contribution of the conventional DEA model is that some of the input and/or output data are characteristically integer- valued. Making use of categorical or ordinal data usually allows including integer-valued data into the analysis can be seen in articles such as Banker and Morey [2]; Kamakura [10] and Rousseau and Semple [18] among authors. The first DEA model allows explicit integrality constraints was developed by Lozano and Villa ([16],[17]). They proposed a mixed linear programming (MILP) DEA model which restricted the computed targets to integers. However, as later argued by Kuosmanen and Kazemi Matin [11] this model does not comply with the minimum extrapolation principle (Banker et.al [1])which is the theoretical foundation of DEA models. Furthermore, the proposed model tends to overestimate the efficiency score. To address these issues, Kuosmanen and Kazemi Matin [14] developed an alternative programming problem with constant return to scale technology (CRS) based upon a new axiomatic

*Corresponding author. aamirteimoori@gmail.com, Tel: +(98)9113330785.
†Department of Mathematics, Islamic Azad University, Rasht Branch, Rasht, Iran.
‡Department of Mathematics, Islamic Azad University, Hadishahr Branch, Hadishahr, Iran.

foundation of natural disposability and natural divisibility for production possibility set involving integers. To be in line with different technologies, the CRS framework was extended to other situations like variable, non-decreasing and non-increasing return to scale environments. Kuosmanen and Kazemi Matin [14] established a new notion of natural convexity, which restricts the feasible convex combination to the subset made up of integer-valued points, but the NDRS variant requires a new postulate of natural augmentability. The connection to NIRS technology is characterized by making use of the earlier natural divisibility axiom. Recently, Kazemi Matin and A. Emrouznejad [11] introduced the notion of boundedness on the subset of output variables in an integer-valued DEA model in an axiomatic approach. Based on the new introduced axiom of "outputs bounded scale," the associated minimal extrapolation PPS is constructed. A mixed integer linear programing (MILP) formulation similar to integer-valued DEA models was suggested for computing output efficiency scores of the units. In a paper presented by D. Khezrimotlagh on DEA conferences in 2015, committed that linear integer DEA models does not mathematically represent the Integer Production Possibility Set (IPPS). Therefore, the benchmarking may not be appropriate. The paper was clearly elucidated this gap and the proposed model can remove the gap. Also, the validity of the method is mathematically proved. A real-life application on university efficiency was examined to depict the differences of several integer DEA models in benchmarking decision making units. G. R. Jahanshahloo and M. Piri [9] proposed a modified model to evaluate the main units in the presence of negative integer data. First, the semi-oriented radial measure (SORM) based on negative data is depicted. After distinguishing the drawbacks of the model, the modified model is introduced, then the model reformed in presence of negative integer values. On the other hand, in recent years, there has been serious attention to modeling undesirable inputs/outputs in DEA literature. Fare et.al [6] developed a non linear DEA model utilizes Farrell -type of efficiency measure to simultaneously increasing of desirable outputs and decreasing the undesirable outputs by the same factor. As an alternative, toward modeling undesirable outputs, one may treat the undesirable outputs as inputs. Fare and Grosskopf [8] argued that this treatment is inconsistent with the physical laws and standard axioms of production theory. The authors imposed an assumption that these undesirable outputs are weakly disposable, also, assumed that all units in the sample apply a uniform abatement factor. Kuosmanen [13] showed how weakly disposable technology can be modeled in the linear structure such that non-uniform abatement factor can be applied. As a major contribution of inclusion of undesirable factors in eco-efficiency measurement is the paper offered by K. Herv Dakpo et.al [5]. The authors acknowledged the new version of the approach by augmenting it with 'interdependence constraints'. In this modeling, there are mainly three options to reduce the levels of detrimental outputs for a fixed technology: firstly, an increase in abatement options through resource diversion (which is accompanied by a reduction of the production of good outputs); secondly, a reduction in pollution-generating inputs (this decreases the levels of intended outputs except for the case of a substitution with non-polluting inputs to maintain the same amount of good outputs production); and thirdly, the use of cleaner inputs, that is to say inputs that generate less bad outputs and maintain at least the same level of good outputs' production. The by-production approach as presented in the paper offers the advantage of disentangling the operational performance and the environmental performance. However, it assumes independence between the two frontiers and thus autonomy of the two performance measures. To overcome this situation, a new modeling approach was proposed by adding additional constraints relative to the pollution-generating inputs. An important concern, that we need to pay much attention in this study, is how to cope with integer-valued undesirable data. As a related contribution, Chen et.al [4] proposed an additive DEA model to deal with integer-valued undesirable output. The proposed additive model compute efficiency scores based on input and output slacks, which provide a clearer view on which variables cause a specific DMU to be inefficient by a certain amount. With these slack results, directions for improvement are easily obtained for each input and output measure. Equipped with proposed set of axioms, we generalize the method to the hybrid case where both real and integer valued inputs and outputs are present. Being able to distinguish weak disposability in current context, the modified Russell measure of efficiency is proposed and a MILP formulation for computing is derived. The specification of weak disposability in this advocated model not only provide decision maker with better insight into the performance of peer DMUs but also help carry out further analysis for managerial decisions. The results are directly applicable to all areas of economics where activity analysis models are employed. The remainder of this study is going to be unfolded as follows: the next section summarizes strategic concepts and previous DEA works, which are clearly related to this study. The weak disposability technology and integer-based DEA models are fully covered in this section. Section 3 describes in detail the conceptual and mathematical framework to measure efficiency under weak disposability assumption for a hybrid case where some of data sets are deemed to be integer while the others are not. An empirical application to a real case is represented in section ??. The contribution is summarized with a conclusion.

2 Preliminaries

2.1 weakly disposable technology

In this section a brief aspect of weak disposability of outputs is introduced. Using the same notation of Kuosmanen [13] the input vector is denoted by $x = (x_1, ..., x_N) \in R_+^N$, desirable or good outputs by $v = (v_1, ..., v_M) \in R_+^M$ and the undesirable or bad outputs by $w = (w_1, ..., w_J) \in R_+^J$. Data for firm $k \in \{1, ...K\}$ is represented by the vector (v^k, w^k, x^k) , the production technology is characterized by production set $Y = \{(v, w, x)|x \in R_+^N$ can produce $(v, w)\}$, or alternatively, by the output set $P(x) = \{(v, w)|(v, w, x) \in Y\}$. Following Shephard [19], weak disposability of outputs is defined as if $(v, w) \in P(x)$ and $0 \leq \theta \leq 1$ then $(\theta v, \theta w) \in P(x)$, $x \in R_+^N$. In case of variable return to scale, the production technology satisfies the following requirements:

A_1) Envelopment: $(v^k, w^k, x^k) \in Y, k \in K$.

A_2) Weak disposability for good and bad outputs:

$$(v, w, x) \in Y, 0 \leq \theta \leq 1 \ \text{ then } \ (\grave{}\text{v}, \grave{}\text{w}, \text{x}) \in \text{Y}.$$

A_3) Free (strong) disposability of inputs and good outputs:

$$(v, w, x) \in Y, (\alpha, \beta) \in R_+^{M+N}, v \geq \beta, \\ \Rightarrow (v - \beta, w, x + \alpha) \in Y$$

A_4) Convexity; Y is closed and convex.

Equipped with these sets of axioms, Fare and Grosskopf [7] have formulated weak-disposable technology in terms of single, scalar valued abatement factor θ as:

$$T_{FG} = \Big\{(v, w, x\ )| \\ \textstyle\sum_{k=1}^K \theta z^k v_m^k \geq v_m \quad m = 1, ..., M \\ \textstyle\sum_{k=1}^K \theta z^k w_j^k = w_j \quad j = 1, ..., J \\ \textstyle\sum_{k=1}^K z^k x_n^k \leq x_n \quad n = 1, ..., N \\ \textstyle\sum_{k=1}^K z^k = 1 \\ z^k \geq 0, \ 0 \leq \theta \leq 1 \quad k = 1, ..., K \Big\} \tag{2.1}$$

Note that the variables $z = (z^1, ..., z^k)$ are referred to intensity weights. To allow for non- uniform abatement factors across firms, Kuosmanen [13] denotes the abatement factor of firm k by θ^k. The author argued the empirical output set as

$$T_K = \{(v, w, x)| \\ \textstyle\sum_{k=1}^K \theta^k z^k v_m^k \geq v_m, m = 1, ..., M, \\ \textstyle\sum_{k=1}^K \theta^k z^k w_j^k = w_j \quad j = 1, ..., J, \\ \textstyle\sum_{k=1}^K z^k x_n^k \leq x_n \quad n = 1, ..., N, \\ \textstyle\sum_{k=1}^K z^k = 1, \\ z^k \geq 0, \ 0 \leq \theta^k \leq 1 \ k = 1, ..., K\} \tag{2.2}$$

Note that formulation (2.1) is a constrained case of formulation (2.2) imposing $\theta^1 = \theta^2 = ... = \theta^k = \theta$.

In particular, the non linear above technology can be restated in an equivalent linear form with a simple substitution of Kuosmanen [13]

$$z^k = \lambda^k + \mu^k$$

Then we must have

$$\lambda^k = \theta^k z^k, \quad \mu^k = (1-\theta^k) z^k$$

Rearranging the terms, the activity analysis technology (2.2) can be rewritten as:

$$\begin{aligned} &T_K^{(L)} = \{(v,w,x)| \\ &\textstyle\sum_{k=1}^{K} \lambda^k v_m^k \geq v_m, \quad m = 1,...,M \\ &\textstyle\sum_{k=1}^{K} \lambda^k w_j^k = w_j, \quad j = 1,...,J \\ &\textstyle\sum_{k=1}^{K} (\lambda^k + \mu^k)\, x_n^k \leq x_n, \ n = 1,...,N \\ &\textstyle\sum_{k=1}^{K} (\lambda^k + \mu^k) = 1 \\ &\lambda^k, \mu^k \geq 0 \ \ k = 1,...,K\} \end{aligned} \tag{2.3}$$

The above formulation (2.3) is now a linear form and the right hand sides of the envelopment constraints are faced up with scaling variables.

2.2 Integer-Valued DEA

In conventional DEA, each observed data is presented by a pair of non negative input and output vector $(X^k, V^k) \in R_+^{N+M}, k \in \{1,...,K\}$. Lozano and Villa [16] assumed that some observed data of DMUs are integer and they partitioned the set of input variables as $I = I^I \bigcup I^{NI}$ and the set of output variables as $V = V^I \bigcup V^{NI}$, where I^I and V^I are the subsets of the corresponding dimensions that must be integer while the others are not. Subsets I^I and I^{NI} as well as V^I and V^{NI} are mutually disjoint also $|I^I|= p \leq N$ and $|V^I|= q \leq M$. The authors proposed the following possibility set:

$$\begin{aligned} T = \Big\{ &(\hat{x},\hat{v})|\, \hat{x}_k \geq \textstyle\sum_{k=1}^{K} z^k x_i^k \quad \forall i, \\ &\hat{v}_k \leq \textstyle\sum_{k=1}^{K} \lambda^k v_r^k \quad ,\forall r, \\ &\hat{x}_k, \hat{v}_k \in Z^{\geq 0}, \ \forall i \in I^I, \forall r \in V^I \Big\} \end{aligned}$$

Kuosmanen and Kazemi- Matin ([14], [15]) proposed the axioms for the scope of integer-valued input-output variables as:

B_1) Natural disposability:

$$(x,v) \in T, \ \ (\alpha,\beta) \in Z_+^{N+M}, y \geq \beta \Rightarrow (x+\alpha, v-\beta) \in T.$$

B_2) Natural divisibility:

$$(x,v) \in T, \ \exists\lambda \in [0,1], \ (\lambda x, \lambda v) \in Z_+^{N+S} \Rightarrow (\lambda x, \lambda v) \in T.$$

B_3) Natural Convexity:

$$(x,v), (x',v') \in Y \Rightarrow$$

$$(\tilde{x}.\tilde{v}) = \lambda(x,v) + (1-\lambda)(x',v')$$

$$, \ 0 \leq \lambda \leq 1 \ (\tilde{x},\tilde{v}) \in Z_+^{N+M} \Rightarrow (\tilde{x},\tilde{v}) \in Y$$

B_4) Integrality:$(x,y) \in T \Rightarrow (x,y) \in Z_+^{N+S}$

B_5) Natural augment ability:

$$(x,v) \in T, \exists\lambda \geq 1, \ \ (\lambda x, \lambda v) \in Z_+^{N+M} \Rightarrow (\lambda x, \lambda v) \in T.$$

B_6)Minimum Extrapolation: T is the intersection of all sets satisfying$(B1)-(B5)$.

Based on the mentioned notations a hybrid setting involving both real and integer valued data set offers a generalized frameworks, which refers to hybrid integer DEA (HIDEA) model. The axiomatic foundation starting from free disposability(A_2) and convexity $(A_4)$of real valued variables(I^{NI}, V^{NI}) and corresponding axioms of natural disposability(B_1) and natural convexity (B_3) of integer-valued variables, also axiom (A_1) is jointly satisfied by all observed data, the following reference technology for the VRS case may be restated as:

$$T_{VRS}^{HIDEA} = \left\{ \begin{array}{l} \begin{pmatrix} x^I & v^I \\ x^{NI} & v^{NI} \end{pmatrix}; (x^I, y^I) \in Z_+^{p+q}; \\ \begin{pmatrix} x^I \\ x^{NI} \end{pmatrix} \geq \sum_{k=1}^{K} z^k \begin{pmatrix} x_k^I \\ x_k^{NI} \end{pmatrix}; \\ \begin{pmatrix} v^I \\ v^{NI} \end{pmatrix} \leq \sum_{k=1}^{K} z^k \begin{pmatrix} v_k^I \\ v_k^{NI} \end{pmatrix}; \\ \sum_{k=1}^{K} z^k = 1, \ z^k \geq 0 \ \forall k \end{array} \right\} \tag{2.4}$$

Since this set is a discrete set of disconnected point, to measure the radial fashion efficiency based on PPS $(T_{VRS}^{HIDEA})$the following mixed integer linear programming (MILP) in input-oriented format proposed by Kuosmanen and Kazemi-matin ([14], [15]) as :

$$\begin{aligned}
&\text{Min } ` - " (\textstyle\sum_{r=1}^{S} s_r^+ + \sum_{n=1}^{N} s_n^- + \sum_{n=1}^{p} s_n^I) \\
&s.t \\
&v_r^o + s_r^+ = \textstyle\sum_{k=1}^{K} z^k v_r^k, \quad r \in V \\
&\theta x_n^o - s_n^- = \textstyle\sum_{k=1}^{K} z^k x_n^k, \quad n \in I^{NI} \\
&\tilde{x}_n - s_n^- = \textstyle\sum_{k=1}^{K} z^k x_n^k, \quad n \in I^I \\
&\theta x_n^o - s_n^I = \tilde{x}_n \qquad n \in I^I \\
&\tilde{x}_n \in Z_+^p \qquad n \in I^I \\
&z^k \geq 0, \qquad k = 1, ..., K \\
&s_r^+ \geq 0, \\
&\forall r \in O,\ s_n^- \geq 0,\ \forall n \in I,\ s_n^I \geq 0,\ \forall n \in I^I.
\end{aligned} \tag{2.5}$$

Symbol ε denote a non-Archimedean infinitesimal and variables s_r^+, s_n^- and s_n^I represent the non-radial slacks and $\tilde{x}_n \in Z_+^p$ is the integer-valued reference points for inputs I^I. DMU_o is efficient if the optimal value of θ equals one. It is worth noting that (HIDEA) model above distinguishes between two input slacks. The first type denoted by $s_n^-(n \in I^I)$ and $s_n^-(n \in I^{NI})$ represents the absolute differences between the convex combination $\sum_{k=1}^{K} z^k x_n^k$(or $\sum_{k=1}^{K} z^k x_n^k$) and the reference points θx_n^o(or $\tilde{x}_n$), while the second type $s_n^I(n \in I^I)$ represents the absolute differences between the reference point $\tilde{x}_n$ and the projection θx_n^o for integer -restricted inputs.

3 Weak Disposability with Integer-valued DEA

The purpose of this section is to show how weakly disposable technology can be modeled whenever some inputs and outputs are restricted to be integer. Based on preceding notations, each feasible activity -which was characterized by a triple non negative input and output vector (v^k, w^k, x^k) $k = 1, ..., K$ can be rewritten as $x = (x^I, x^{NI})$ $v = (v^I, v^{NI})$ and $w = (w^I, w^{NI})$.

Without a less of generality, $x^I(I \in I^I)$, $v^I(I \in V^I)$ and $w^I(I \in W^I)$are the dimensions satisfies integrality assumptions (B_4). Suppose that $|I^I| = p \leq N$,$|V^I| = q \leq M$ and $|W^I| = \alpha \leq J$. In order to deal with weak disposability in a systematic fashion in hybrid setting- involving both real and integer valued data- a revised set of axioms is needed. Commencing from free disposability of inputs and good outputs (A_3), weak disposability of good and bad outputs (A_2) and convexity (A_4) for real-valued data (x^{NI}, v^{NI}, w^{NI}), the corresponding axioms of integer-valued variables are proposed as:

B_1) Natural disposability of inputs and good outputs:
$(x, v, w) \in T, (\alpha,\beta) \in Z_+^{N+M}, v \geq \beta \Rightarrow (x + \alpha, v - \beta, w) \in T.$

B_2) Integer weak disposability for good and bad outputs:
$(x, v, w) \in T,\ \exists\theta \in [0, 1], (x, \theta v, \theta w) \in Z_+^{M+J} \Rightarrow (x, \theta v, \theta w) \in T.$

B_3) Natural convexity:
$(x, v, w), (x', v', w') \in T(\tilde{x}, \tilde{v}, \tilde{w}) = \lambda(x, v, w) + (1 - \lambda)(x', v', w'),\ 0 \leq \lambda \leq 1, (\tilde{x}, \tilde{v}, \tilde{w}) \in Z^+ \Rightarrow (\tilde{x}, \tilde{v}, \tilde{w}) \in T.$

B_4) Natural divisibility:
$(x, v, w) \in T,\ \exists\lambda \in [0, 1], (\lambda x, \lambda v, \lambda w) \in Z_+^{N+M+J} \Rightarrow (\lambda x, \lambda v, \lambda w) \in T$

B_5) Natural augment ability:
$(x, v, w) \in T, \exists\lambda \geq 1, (\lambda x, \lambda v, \lambda w) \in Z_+^{N+M+J} \Rightarrow (\lambda x, \lambda v, \lambda w) \in T$

The axioms of natural disposability (B_1), natural divisibility (B_4), natural convexity (B_3) and natural augment ability (B_5) have been introduced in Kuosmanen and Kazemi-Matin ([14], [15]). The notation of integer weak disposability (B_2') is a substantial format of weak disposability axioms (A_2) .This axiom states that the abatement of integer-valued undesirable outputs takes place through scaling down of the activity level if it results an integer valued output vector. Rather, natural divisibility (B_4) and natural augment- ability (B_5) postulate that a division or augmentation of a production plan is possible. Also, these axioms can be seen as alternative variants of return to scale axioms. Equipped with these sets of assumptions, the hybrid integer DEA (HIDEA) reference technology for the VRS case can be stated as:

$$T_{VRS}^{HIDEA} = \left\{ \begin{array}{l} \begin{pmatrix} x^I & v^I & w^I \\ x^{NI} & v^{NI} & w^{NI} \end{pmatrix}; \\ (x^I, v^I, w^I) \in Z_+^{N+M+J}, \\ x^I \geq \sum_{k=1}^{K} z^k x_n^k, n \in I^I, \\ x^{NI} \geq \sum_{k=1}^{K} z^k x_n^k, \quad n \in I^{NI}, \\ v^I \leq \sum_{k=1}^{K} \theta^k z^k v_m^k, \quad m \in V^I \\ v^{NI} \leq \sum_{k=1}^{K} \theta^k z^k v_m^k, \ m \in V^{NI} \\ w^I = \sum_{k=1}^{K} \theta^k z^k w_j^k, \quad j \in W^I \\ w^{NI} = \sum_{k=1}^{K} \theta^k z^k w_j^k, \ j \in W^{NI} \\ \sum_{k=1}^{K} z^k = 1 \\ 0 \leq \theta^k \leq 1 \\ z^k \geq 0 \end{array} \right\} \quad (3.6)$$

The following theorem establishes the axiomatic foundation of this reference technology under revised sets of assumptions to the subsets (I^{NI}, V^{NI}, W^{NI}) and (I^I, V^I, W^I) respectively.

Theorem 3.1 *Production set T_{VRS}^{HIDEA} is the minimum extrapolation production possibility set if subsets (I^{NI}, V^{NI}, W^{NI}) Satisfy axioms $(A_2 - A_4)$ and subsets (I^I, V^I, W^I) satisfy natural disposability of inputs and good outputs (B_1), integer weak disposability of good and bad outputs (B_2') and natural convexity (B_3) also axiom (A_1) is jointly satisfied by all observed data set.*

Proof. It suffices to show that the axioms (B_1), (B_2') and (B_3) are simply integrality restricted cases of (A_3), (A_2) and(A_4). Moreover, the same intensity variable z apply to both subsets (I^{NI}, V^{NI}, W^{NI}) and (I^I, V^I, W^I). The minimum extrapolation theorem for the case of real-valued data (I^{NI}, V^{NI}, W^{NI}) with axioms $(A_1 - A_4)$ has been formally proved by Banker et.al [1], the case of (I^I, V^I, W^I) with axioms B_1, B_2' and B_3 was proved in Kuosmanen and Kazemi-Matin[15].

In spirit of alternative return to scale specification in IDEA framework, the axioms of natural divisibility (B_4) and natural augment -ability (B_5) along with other postulates can interpret NIRS and NDRS variants of IDEA technology respectively. Moreover, the multiplier θ^k used in this VRS technology enables the reduction of the level of bad outputs if accompanied by the reduction of desirable outputs in the same proportion across all firms. To linearize reference technology T_{VRS}^{HIDEA}, the intensity weight of firm k can be partitioned into two components $z^k = \lambda^k + \mu^k$. Using this substitution from Kuosmanen [13] the production technology (3.6) converts into the following linear form:

$$T_{VRS}^{HIDEA} = \left\{ \begin{array}{l} \begin{pmatrix} x^I & v^I & w^I \\ x^{NI} & v^{NI} & w^{NI} \end{pmatrix}; \\ (x^I, v^I, w^I) \in Z_+^{N+M+J}, \\ x^I \geq \sum_{k=1}^{K} (\lambda^k + \mu^K) x_n^k, \quad n \in I^I \\ x^{NI} \geq \sum_{k=1}^{K} (\lambda^k + \mu^K) x_n^k, \ n \in I^{NI} \\ v^I \leq \sum_{k=1}^{K} \lambda^k v_m^k, \quad m \in V^I \\ v^{NI} \leq \sum_{k=1}^{K} \lambda^k v_m^k, \quad m \in V^{NI} \\ w^I = \sum_{k=1}^{K} \lambda^k w_j^k, \quad j \in W^I \\ w^{NI} = \sum_{k=1}^{K} \lambda^k w_j^k, \quad j \in W^{NI} \\ \sum_{k=1}^{K} (\lambda^k + \mu^K) = 1 \\ \lambda^k, \mu^k \geq 0 \end{array} \right\} \quad (3.7)$$

The technology above is a non- monotonic and non-convex set of disconnected points. As Kuosmanen and Kazemi-matin([14], [15]) noted, a modified Farrell input efficiency measure is needed. Applying this set, the modified efficiency scores relative to the general T_{VRS}^{HIDEA} reference technology can be formulated by solving the following mixed integer linear programming (MILP) problem:

$$
\begin{array}{ll}
\sigma = Min\frac{1}{N+J}\left[\sum_{n=1}^{N}\theta_n + \sum_{j=1}^{J}\varphi_j\right] & \\
s.t & \\
\sum_{k=1}^{K}(\lambda^k+\mu^k)x_n^k = \tilde{x}_n - s_n^-, & n\in I^I \\
\tilde{x}_n = \theta_n x_n^o - s_n^I, & n\in I^I \\
\sum_{k=1}^{K}(\lambda^k+\mu^k)x_n^k = \theta_n x_n^o - s_n^-, & n\in I^{NI} \\
\sum_{k=1}^{K}\lambda^k v_m^k = \tilde{v}_m + s_m^+, & m\in V^I \\
\tilde{v}_m = v_m^o + s_m^I, & m\in V^I \\
\sum_{k=1}^{K}\lambda^k v_m^k = v_m^o + s_m^+, & m\in V^{NI} \\
\sum_{k=1}^{K}\lambda^k w_j^k = \tilde{w}_j, & j\in W^I \\
\tilde{w}_j = \varphi_j w_{jo}^I - s_j^I, & j\in W^I \\
\sum_{k=1}^{K}\lambda^k w_j^k = \varphi_j w_{jo}^O, & j\in W^{NI} \\
\sum_{k=1}^{K}(\lambda^k+\mu^k) = 1 & \\
\lambda^k,\mu^k \geq 0 & \\
\tilde{x}_n,\tilde{v}_m,\tilde{w}_j \in Z_+ & \\
\varphi_j \geq 1 \quad j=1,...,J & \\
0\leq\theta_n\leq 1 \quad n=1,...,N & \\
s_n^-, s_m^+ \geq 0 & \\
s_n^I, s_m^I, s_j^I \geq 0,\ n\in I^I, m\in V^I, j\in W^I. &
\end{array} \tag{3.8}
$$

The MILP formulation above measures the efficiency score of firm 0 in terms of the abatement potential factor in integer and real valued inputs and outputs. Also, the constraint $\varphi_j \geq 1 \;\; j=1,...,J$ and $0\leq\theta_n\leq 1 \quad n=1,...,N$ are the requirements for dominance. It is worth to note both data set exhibit that variable return to scale (VRS) and the objective function can be represented as $\frac{1}{N+J}\left[\sum_{n=1}^{N}\theta_n + \sum_{j=1}^{J}\mu_j\right]$, the Russell-input and bad output measure of efficiency. So, the optimal value of model (3.8) is equal to the Russell efficiency measure defined with respect to the T_{VRS}^{HIDEA} reference technology. In essence, model (3.8) and its constraints imposed both integer and real restrictions on inputs and undesirable outputs whilst weak disposability influences the output set. In terms of efficiency measurement, we scope on minimizing the potential changes in inputs and undesirable outputs. Top of all in the context of discrete set of points, the evaluated DMU can be projected optimally close to the non -negative integer feasible point through solving MILP formulation above.

4 Application

In order to see how weak disposability influences the hybrid output set, let us consider sixteen decision making units with three inputs which is denoted by x_1 and $x_2$as integer valued and $x_3$as real-valued and four outputs. The last output component w_1 is characterized as undesirable one. w_1 Is represented integer-values. The other componentsv_1 and v_2 are considered to be the number of desirable production. The first is integer and the last indicates real valued. Table (1) summarizes all data set which was taken from Chen et.al [4]. In order to shed a light on proposed approach, the modified additive model introduced by Chen et.al [4] is recorded here. Again assume that, $|I^I|= p \leq N$, $|V^I|= q \leq M$ and $|W^I|= \alpha \leq J$. Also, according to the attributes the input vector $x = (x^I, x^{NI})$ is categorized into desirable and undesirable integer-valued and real-valued. The notation $x_n^G, x_n^B (n \in I^{NI})$ presents the real-valued input quantities and $x_n^G, x_n^B (n \in I^I)$ depicts integer-valued desirable and undesirable inputs. Also the number of desirable and undesirable integer-valued input sets can be shown as the index p_1 and p_2 with $p = p_1 + p_2$. Based on preceding notations, the model has the format as follows:

$$
\begin{array}{ll}
\rho = Max\ \frac{1}{(N-p)+p_1+p_2+q+(M-q)+\alpha+(J-\alpha)} & \\
\times\left[\sum \frac{s_n^{oG-}}{x_n^{Go}} + \frac{s_n^{oB-}}{x_n^{Bo}} + \frac{s_n^{oGI-}}{x_n^{Go}} + \frac{s_n^{oBI-}}{x_n^{Bo}}\right. & \\
\left. + \frac{s_m^{oI+}}{v_m^o} + \frac{s_m^{oNI+}}{v_m^o} + \frac{s_j^{oI+}}{w_j^o} + \frac{s_j^{oNI+}}{w_j^o}\right] & \\
s.t. & \\
\sum_{k=1}^{K}\lambda^k x_n^{Gk} = x_n^{Go} - s_n^{oG-}, & n\in I^{NI} \\
\sum_{k=1}^{K}\lambda^k x_n^{Bk} = x_n^{Bo} + s_n^{oB-}, & n\in I^{NI} \\
\sum_{k=1}^{K}\lambda^k x_n^{Gk} \leq x_n^{Go} - s_n^{oGI-}, & n\in I^{I} \\
\sum_{k=1}^{K}\lambda^k x_n^{Bk} \geq x_n^{Bo} + s_n^{oBI-}, & n\in I^{I} \\
\sum_{k=1}^{K}\lambda^k v_m^k \geq v_m^o + s_m^{oI+}, & m\in V^I \\
\sum_{k=1}^{K}\lambda^k v_m^k = v_m^o + s_m^{oNI+}, & m\in V^{NI} \\
\sum_{k=1}^{K}\lambda^k w_j^k \leq w_j^o - s_j^{oI+}, & j\in W^I \\
\sum_{k=1}^{K}\lambda^k w_j^k = w_j^o - s_j^{oNI+}, & j\in W^{NI} \\
\sum_{k=1}^{K}\lambda^k = 1 & \\
\lambda^k \geq 0, \quad k=1,...,K, & \\
s_n^{oG-}, s_n^{oB-}, s_m^{oNI+}, s_j^{oNI+} \geq 0 & \\
s_n^{oGI-} \in Z_+^{P_1}, s_n^{oBI-} \in Z_+^{P_2}, & \\
s_m^{oI+} \in Z_+^{q}, s_j^{oI+} \in Z_+^{\alpha} &
\end{array} \tag{4.9}
$$

Table 1: Efficiency Score and Dominance Factors.

DMU	x_1Int	x_2Int	x_2 Cont	v_1 Cont	v_2 Int	w_1Int
1	403	1120	5331608	18363085	31277605	154
2	411	1125	5883209	19695025	30840680	127
3	420	1112	5940485	19028804	30586165	152
4	452	1121	6626290	20103803	33175128	140
5	495	1126	8154603	21189140	36092750	194
6	473	1146	8502509	21189140	37271623	175
7	472	1087	9031125	21136856	38828087	161
8	454	1132	8788094	20801147	309107870	142
9	432	1108	9157097	21698043	38143325	180
10	427	1045	8868206	21051481	33723130	171
11	438	979	8398829	20114530	32698925	159
12	450	923	9083108	22023502	32641359	232
13	435	851	7470562	17400575	29972388	150
14	424	869	7188360	17101044	28763740	147
15	420	829	6665653	16937531	25005947	118
16	474	879	6489033	17681062	22596922	123

In the above model, s_n^{oG-} , s_n^{oB-} , s_m^{oNI+}, s_j^{oNI+} and

$$s_n^{oGI-} \in Z_+^{P_1}$$

,

$$s_n^{oBI-} \in Z_+^{P_2}$$

,

$$s_m^{oI+} \in Z_+^{q}$$

,

$$s_j^{oI+} \in Z_+^{\alpha}$$

are non-radial slack vectors of inputs and outputs of under evaluated DMU. Although the optimal value of the model does not depend on the units of measurement in inputs and outputs. Top of all, this additive model provides a closer on which variables cause a specific DMU to be inefficient by a certain amount. With these slack results, directions for improvements are easily obtained for each input and output measure. It is worth noting that the inequality is used in model (4.9) for integer-restricted inputs and outputs because the convex combinations for frontier DMUs are not necessarily integer-valued.

Therefore, $x_n^{Go} - s_n^{oGI-}(n \in I^I)$, $x_n^{Bo} + s_n^{oBI-}(n \in I^I)$, $v_m^o + s_m^{oI+}(m \in V^I)$ and $w_j^o - s_j^o, (j \in W^I)$, the reference target for integer factors may or may not be equal to the their projections on the efficient frontier, but must be dominated by their convex combinations of frontier DMUs. To obtain an efficiency score between zero and one, let the optimal solution of the model is $(\lambda^{k*}, s_n^{*oG-}, s_n^{*oB-}, s_m^{*oNI+}, s_j^{*oNI+}, s_n^{*oGI-},$ $s_n^{*oBI-}, s_m^{*oI+}, s_j^{*oI+})$, the efficiency score of DMU_o can be computed as below:

$$\rho_o^* = \frac{1 - \frac{\left(\sum \frac{s_n^{*oG-}}{x_n^{Go}} + \frac{s_n^{*oGI-}}{x_n^{Go}} + \frac{s_j^{*oI+}}{w_j^o} + \frac{s_j^{*oNI+}}{w_j^o}\right)}{(N-p)+p_1+(J-\alpha)+\alpha}}{1 + \frac{\left(\sum \frac{s_n^{*oB-}}{x_n^{Bo}} + \frac{s_n^{*oBI-}}{x_n^{Bo}} + \frac{s_m^{*oI+}}{v_m^o} + \frac{s_m^{*oNI+}}{v_m^o}\right)}{(M-q)+q+(N-p)+p_2}}$$

It is true that the value of ρ_o^* falls between zero and one. A larger efficiency value indicates that the DMU is closer to the efficient frontier and has a better performance. Also the unit is efficient if and only if $\rho_o^* = 1$. In other words, all slacks in model (4.9) are zero. The results of employing

Table 2: Efficiency Score and Dominance Factors.

DMU	ρ_o^*	σ^*	θ_1^*	θ_2^*	θ_3^*	φ_1^*
1	1	1	1	1	1	1
2	1	1	1	1	1	1
3	0.9397	0.98	0.9713	1	0.9558	1
4	1	0.97	0.9207	0.9921	0.9847	1
5	0.9132	0.94	0.8816	0.8853	0.9758	1
6	0.9415	0.96	0.9087	0.9393	0.9661	1
7	1	0.97	0.9081	1	0.9044	1.0630
8	1	1	1	1	1	1
9	1	1	1	1	1	1
10	1	1	1	1	1	1
11	1	0.98	0.9809	1	0.8828	1.0656
12	1	1	1	1	1	1
13	1	0.97	0.9698	1	0.9109	1
14	1	0.97	0.9881	0.9860	0.9151	1
15	1	1	1	1	1	1
16	1	1	1	1	1	1

model (4.9) are presented in the second column of Table 2. Applying linear program (3.8), the obtained efficiency score is presented in Table 2. When weak disposability is introduced, employing the model based on technology (3.7) presents the following scores. The first column of Table 2 reports the efficiency score of the additive model (4.9) . Compared with the thirteen efficient units which were reported in model (4.9), the proposed approach records only eight efficient DMUs. For this application, the additive model (4.9) is not too helpful. Because model (4.9) considers undesirable factors behaving like inputs. Therefore, in this computation effort through applying weak disposable assumption for undesirable outputs the number of efficient unit decreases. It is worth to stress that the dominance factor which indicates weak disposability of undesirable outputs and inputs play notable role in these results. The rest columns of Table 2 report the optimal values for non uniform dominance factors across all firms. As regards, the proposed formulations can be fallen beyond the scope of hybrid return to scale environments in efficiency measurement contexts.

5 Conclusion

In many real practical problems it is not uncommon to encounter production outputs that are undesirable or can only take integer-values. Standard DEA models assume that all inputs and/or outputs are continuous and real-valued; therefore, the identified targets are likely to be fractional and not applicable in practice. To tackle with integer constraints and undesirable outputs different type of integer-valued efficiency measure were proposed to calculate the performance of all DMUs. Furthermore, the correct specification of weak-disposability in two types of models was important not only for productivity and efficiency measurment but also for estimating the dominance possibility. As confirmed by the numerical application the weak disposability property for both real and integer valued data set can be seen as abatement achieved through decrease in the activity level. Finally, all the discussion can be applied through hybrid return to scale (RTS) situations.

References

[1] R. J. Banker, A. Charnes, W. W. Cooper, Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science* 9 (1984) 1078-1092.

[2] R. J. Banker, R. C. Morey, The use of categorical variables in data envelopment analysis, *Management Science* 32 (1986) 1613-1627.

[3] A. Charnes, W. W. Cooper, E. Rhodes, Measuring the efficiency of decision making units, *European Journal of Operation Research* 2 (1978) 429-444.

[4] Chien-M. Chen, J. Du, J. Huo, J. Zhu, Undesirable factors in integer-valued DEA: Evaluating the operational efficiencies of city bys systems considering safety systems, *Decision Support Systems* 54 (2012) 330-335.

[5] K. H. Dakpo, P. Jeanneaux, L. Latruffe, Inclusion of undesirable factors in production technology modeling: The case of greenhouse gas emissions in French meat sheep farming, *Working Paper SMART-LERECO* 14 (2014) 1-47.

[6] R. Fare, S. Grosskopf, K. Lovell, C. Pasurka, Multilateral productivity comparisons when some outputs are undesirable: a non parametric approach, *The review of Economics and Statistics* 71 (1989) 90-98.

[7] R. Fare, S. Grosskopf, Non parametric Productivity Analysis with Undesirable outputs: Comment, *American Journal of Agriculture Economics* 85 (2003) 1070-1074.

[8] R. Fare, S. Grosskopf, New Directions: Efficiency and Productivity, Kluwer Academic Publishers, *Boston*, (2004).

[9] G.R. Jahanshahloo, M. Piri, Data Envelopment Analysis (DEA) with integer and negative inputs and outputs, *Data Envelopment Analysis and Decisions Science* (2013) 1-15.

[10] W. A. Kamuraka, A note on the use of categorical variables in data envelopment analysis, *Management Science* 34 (1988) 1273-1276.

[11] R. Kazemi-matin, R. Emrouznejad, An Integer-valued data envelopment analysis model with bounded outputs, *International Transactions in Operational Research* 18 (2011) 741-749.

[12] D. Khezrimotlagh, Benchmarking decision making units with integer values in Data Envelopment Analysis, *International Conferences on DEA in Economics and Finance, Technical University of Ostrava* (2015).

[13] T. Kuosmanen, Weak disposability in nonparametric analysis with undesirable outputs, *American Journal of Agriculture Economics* 87 (2005) 1077-1082.

[14] T. Kuosmanen, R. Kazemi-matin, Theory of integer-valued data envelopment analysis, *European Journal of Operation Research* 218 (2009) 186-192.

[15] T. Kuosmanen, R. Kazemi-matin, Theory of integer-valued data envelopment analysis under alternative returns to scale axioms, *Omega* 37 (2009) 988-995.

[16] S. Lozano, G. Villa, Data Envelopment Analysis for integer-valued inputs and outputs, *Computers and Operations Research* 33 (2006) 3004-3014.

[17] S. Lozano, G. Villa, Integer DEA models. In: Zhu, J., Cook, W. D., Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis, *Springer, New York,* (2007) 271-290.

[18] J. J. Rousseau, J. H. Semple, Categorical outputs in data envelopment analysis, *Management Science* 39 (1993) 384-386.

[19] R. W. Shephard, Theory of cost and production functions, *Princeton University Press* (1970).

[20] K. Tone, A slacks-based measure of efficiency in data envelopment analysis, *European Journal of Operational Research* 130 (2001) 498-509.

Alireza Amirteimoori is a professor in Applied Mathematics Operations Research group in Islamic Azad University in Rasht, Iran. His research interests lie in the broad area of performance management with special emphasis on the quantitative methods of performance measurement, and especially those based on the broad set of methods known as data envelopment analysis (DEA). Amirteimoori's papers appear in journals such as International Journal of Mathematics in Operations Research, IMA Journal of Management Mathematics, Applied Mathematics and Computation, Journal of the Operations Research Society of Japan, Journal of Applied Mathematics, Expert Systems, Journal of Global Optimization, Decision Support Systems, Optimization, Central European Journal of Operations Research, Expert Systems with Applications, Transportation Research Part D: Transport and Environment, International Journal of Advanced Manufacturing Technology, RAIRO-Operations Research, Applied Mathematical Letters, International Journal of Production Economics, Measurement and the like.

Mahnaz Maghbouli is an Assistance Professor of Applied Mathematics and Operation Research at Islamic Azad University in Iran (Hadishahr Branch). Her research interests lie in the broad area of performance management and multi-criteria decision making with special emphasis on the quantitative methods of performance measurement.