



Munich Personal RePEc Archive

Weighted Additive DEA Models Associated with Dataset Standardization Techniques

Chen, Kaihua

February 2014

Online at <https://mpra.ub.uni-muenchen.de/55072/>
MPRA Paper No. 55072, posted 06 Apr 2014 16:28 UTC

Weighted Additive DEA Models Associated with Dataset Standardization Techniques

Kaihua Chen¹

Institution of Policy and Management, Chinese Academy of Sciences
No. 15 ZhongGuanCunBeiYitiao, Alley, Beijing 100190, People's Republic of China
Tel/Fax:+86-10-59358707; E-mail: chenkaihua@casipm.ac.cn

Mingting Kou

School of management, University of Chinese Academy of Sciences
No. 80 ZhongGuanCunDong, Road, Beijing 100190, People's Republic of China
Tel/Fax:+86-10-59358707; E-mail: koumingting@ucas.ac.cn

Abstract This paper uncovers the “mysterious veil” above the formulations and concerned properties of existing weighted additive data envelopment analysis (WADD) models associated with dataset standardization techniques. Based on the truth that the formulation of objective functions in WADD models seems random and confused for users, the study investigates the correspondence relationship between the formulation of objective functions by statistical data-based weights aggregating slacks in WADD models and the pre-standardization of original datasets before using the traditional ADD model in terms of satisfying unit and translation invariance. Our work presents a statistical background for WADD models' formulations, and makes them become more interpretive and more convenient to be computed and practically applied. Based on the pre-standardization techniques, two new WADD models satisfying unit invariance are formulated to enrich the family of WADD models. We compare all WADD models in some concerned properties, and give special attention to the (in)efficiency discrimination power of them. Moreover, some suggestions guiding theoretical and practical applications of WADD models are discussed.

Keywords: Data envelopment analysis; Weighted additive models; Formulations and applications; Dataset standardization techniques

¹ Corresponding Author.
Tel/Fax:+86-10-59358707.
E-mail: chenkaihua@casipm.ac.cn (Kaihua C.)

1. Introduction

In the complex evaluation system with multiple variables, there are usually differences in both the orders of magnitude (scale) within a dataset and the units of measurement (dimension) across datasets. It's not advisable and appropriate to directly use original observation data as the entry values of evaluation models in the evaluation systems with "multi-unit" and "disparate orders of magnitude". A dimensionless processing of original datasets, which doesn't damage the measure of performance results, is usually necessary before calculating models.

Through dimensionless preprocessing of original observation datasets, the evaluation models will escape the harms resulting from the differences in orders of magnitude and units of measurement, which makes measure indicators be commensurable and play in the fair field. If the evaluation models (or formulas) can overcome unit heterogeneity by themselves without a dimensionless process, they are unit invariant in terminology (Lovell and Pastor, 1995). For example, the standard radial data envelopment analysis (DEA) models including BCC (Banker et al., 1984) and CCR (Charnes et al., 1978) are unit invariant. They needn't pre-standardize original datasets for obtaining comparable efficiency results when using them. However, among non-radial DEA models, the additive DEA (ADD) model (Charnes et al., 1985) doesn't satisfy unit invariance, whose result is nothing but a sum of heterogeneous (unit and scale variety) and incommensurable slacks. Evidently, it's necessary to pre-standardize original datasets when applying it in the multivariable evaluation systems with "multi-unit" and "disparate orders of magnitude".

Besides, the basic DEA models are usually not compatible with negative and zero numbers, usually being "nail-biting". However, for many practical decisions, zero and negative numbers really exist in observation data sets (Pastor and Ruiz, 2007). Negative numbers can be used to measure lost profits amount and cost savings amount. Zero can be used to measure no profits and no cost savings or no inputs and outputs. A common and feasible method coping with those cases is to add a positive constant number to the dataset with negative and zero numbers translating it into a positive dataset. If the translation doesn't change the calculation result, then the evaluation model (or formula) is translation invariant in terminology (Pastor, 1996). For example, the ADD model and the constraint formulas on outputs (inputs) in input-(output-) oriented BCC model are translation invariant.

The properties of unit invariance and translation invariance in DEA models have been

argued in many previous literatures (Ali and Seiford, 1990; Pastor, 1994; Lovell and Pastor, 1995; Thrall, 1996; Pastor, 1996; Cooper et al., 1999; Cooper et al., 2007; Pastor and Ruiz, 2007), which recently were summarized in Sueyoshi and Sekitani (2009) and Cook and Seiford (2009). The previous literature has investigated the extended ADD models mainly in terms of whether satisfying unit invariance and translation invariance or not, or whether satisfying the sense of efficiency or not with respect to the efficiency values ranging $[0, 1]$ or not, or from the discrimination power of (in)efficiency scores with respect to the distribution of efficiency values within the score set. For example, RAM (Range Adjusted Measure) (Cooper et al., 1999) is a popular extension of the ADD model, which performs well in the first two items. However, the (in)efficiency values by it crowd in a small range of $[0, 1]$, which are approximately unity and short of discrimination power (Sueyoshi and Sekitani, 2009).

Because the constraints for inputs and outputs are unit and translation invariance under variable return to scale, the extending studies about the ADD model are mainly to make the objective functions of the extended ADD models satisfy unit invariance or even translation invariance. Since weights on slacks in objective functions influence main attractive properties of the extended ADD models (Ali et al., 1995)^②, the existing literature about extended ADD models mainly focuses on choosing weights on the slacks in the objective functions. It is done by introducing weights assigned to the inefficiencies (slacks) with which they are associated (Cooper et al., 1999). Then, the weighted ADD (referred to as WADD) models are formulated.

In this situation, the choices of suitable weights become a hot topic for the development of WADD models. Thrall (1996) noted that the choices he considered for weights represent “value judgments” that reflect considerations not embodied in the datasets. This kind of value-judgment-oriented subjective weights is to some degree suitable, and often needed in some management practices (Cooper et al., 1999). However, it seems that it is more desirable and favorable to orient these choices to more “objective” criteria, so that investigators doing scientific research, for instance, will reach the same (or similar) conclusions from the same (or similar) bodies of data (Cooper et al., 1999). In fact, objective weights based on statistical data-based weights are more popular in the researches of existing WADD models, such as

^② Components of efficiency evaluation in data envelopment analysis include (i) the form of envelopment surface, (ii) the orientation, and (iii) the pricing mechanism implicit in the multiplier lower bounds, i.e., weights (Ali et al., 1995). ADD is non-orientated and the manner in which the projection (onto the fixed development surface depending on the constraints) is effected in the weighted ADD model only depends weights.

MIP-WADD (Charnes et al., 1985), SUM-WADD (Ali et al., 1995), NOR-WADD (Lovell and Pastor, 1995) and RIM-WADD (Pastor, 1994) as well as RAM-WADD (Cooper et al., 1999). However, they didn't systematically provide the statistical theory background for choices of those statistical data-based weights, which makes the formulations of WADD models appear random and confused for users. It's not favorable to practical applications and model extensions. For example, some extensions are only to confine the resulting (in)efficiency scores into $[0, 1]$ and satisfy a sense of (in)efficiency or to make the model be translation invariant, but usually bring difficulties to the practical applications. So it's valuable to systematically investigate the statistical background and the motivation of choosing weights in formulating WADD models.

In this study, we, confining ourselves to "statistical data-based weights", focus on investigating and extending WADD models from a practical and statistical perspective. A systematic investigation of the weighted formulations of objective functions in existing WADD models is implemented to reveal the inherent consistency relationship between the pre-standardization of original datasets using the ADD model and the formulations of objective functions in the WADD models. The finding not only induces us to develop two new WADD models based on two non-referred standardization techniques, but also provides us a simple transformation route to deal with the complex WADD models. Moreover, a modeling application of the WADD models is introduced which improves the multi-stage models proposed by Fried et al. (1999, 2002) and makes them unit and translation invariant.

The remaining structure of this paper is organized as follows: Section 2 summarizes the standardization techniques of the preprocessing of original datasets. Section 3 investigates the unit invariance and translation invariance of existing WADD models as well as the inefficiency discrimination power of them. Section 4 proposes two new weighted additive models for inefficiency measurement by assigning maximum or mean to the slacks of objective function as weights. Section 5 is focused on the empirical comparison of WADD models in the (in)efficiency discrimination power. Some applications of WADD models are discussed in Section 6, which include an extended application to the multi-stage model proposed by Fried et al. (1999, 2002). Conclusion is presented in Section 7.

2. Standardization techniques for original datasets

There are n decision making units (DMUs) to be evaluated. For each

DMU_{*j*} (*j* = 1, 2, ..., *n*), there are *m* cost (input) variables x_{ij} (*i* = 1, 2, ..., *m*) and *s* profit (output) variables y_{rj} (*r* = 1, 2, ..., *s*) available.

To avoid the damage to comparable evaluation results because of the differences in the orders of magnitude and the units of measurement under the multivariable evaluation environment, original observation datasets are usually standardized by transformation formulas (S1) ~ (S5) as follows. The transformed datasets satisfy “dimensional homogeneity” because the original datasets are scaled by descriptive statistics with the same dimensions as the original datasets, such as standard deviation, mean, maximum, sum and range of original observation datasets. This means that these transformation formulas can make a raw/original in-commensurable dataset transformed into a standardized commensurable dataset.

$$\hat{x}_{ij} = \frac{x_{ij} - \bar{x}_i}{\sigma_i^-}, \quad \hat{y}_{rj} = \frac{y_{rj} - \bar{y}_r}{\sigma_r^+} \quad (\text{S1})$$

$$\hat{x}_{ij} = \frac{x_{ij}}{x_i^U}, \quad \hat{y}_{rj} = \frac{y_{rj}}{y_r^U} \quad (\text{S2})$$

$$\hat{x}_{ij} = \frac{x_{ij} - x_i^L}{x_i^U - x_i^L}, \quad \hat{y}_{rj} = \frac{y_{rj} - y_r^L}{y_r^U - y_r^L} \quad (\text{S3})$$

$$\hat{x}_{ij} = \frac{x_i^U - x_{ij}}{x_i^U - x_i^L}, \quad \hat{y}_{rj} = \frac{y_r^U - y_{rj}}{y_r^U - y_r^L} \quad (\text{S3}^*)$$

$$\hat{x}_{ij} = \frac{x_{ij}}{\bar{x}_i}, \quad \hat{y}_{rj} = \frac{y_{rj}}{\bar{y}_r} \quad (\text{S4})$$

$$\hat{x}_{ij} = \frac{x_{ij}}{\sum_{j=1}^n x_{ij}}, \quad \hat{y}_{rj} = \frac{y_{rj}}{\sum_{j=1}^n y_{rj}} \quad (\text{S5})$$

Here, $r = 1, 2, \dots, s$; $i = 1, 2, \dots, m$; statistics $\sigma_i^- = \sqrt{\sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 / n - 1}$ and $\sigma_r^+ = \sqrt{\sum_{j=1}^n (y_{rj} - \bar{y}_r)^2 / n - 1}$ are the sample standard deviations of the cost (input) and profit (output) variables respectively; statistics $\bar{x}_i = \sum_{j=1}^n x_{ij} / n$ and $\bar{y}_r = \sum_{j=1}^n y_{rj} / n$ are the mean of dataset *i* and dataset *r* respectively; statistics $x_i^U = \max_j(x_{ij})$, $x_i^L = \min_j(x_{ij})$, $y_r^U = \max_j(y_{rj})$ and $y_r^L = \min_j(y_{rj})$.

Among them, (S1) ~ (S3) are most frequently used in the practical applications. The mean and standard deviation of the transformed dataset by (S1) are zero and unity respectively, and (S2) and (S3) make the transformed dataset rang in [0, 1]. (S4) makes the

standardized dataset fluctuate as the center of mean, and (S5) makes the sum of standardized dataset be unity. More important, formulas (S1) and (S3) include a centralization/translation transformation as well as a dimensionless transformation, while formulas (S2), (S4) and (S5) only include a dimensionless transformation. In virtue of dimensionless transformation, (S1) ~ (S5) are unit invariant, which make transformed data apart from the deceptive effect of unit/dimension differences of measurement and order differences of magnitude on the real relative variances. Besides, when an original dataset add or reduce a constant, the transformed dataset by (S1) and (S3) don't vary, i.e., formulas (S1) and (S3) are translation invariant, which is due to that their denominators and numerators are the difference between the observation data and the maximum, minimum or average of observation dataset. These descriptive statistics (including observation data) have the same change, so the difference form eliminates/moderates the influence of simultaneous changes on the denominators and numerators. This means that translation invariance can be obtained by centralization/translation transformation of original datasets.

So it's practically valuable to substitute original input and output data for transformed input and output data in the ADD model, which is consistent with the procedure of preprocessing original datasets by (S1) ~ (S5) before using the basic ADD model. This substitution process can be referred to as the standardization of the ADD model, which is consistent with the formulation of WADD models. The study below shows the inherent consistency between the formulation of WADD models and the pre-standardization of the original datasets on using the ADD model.

Table 1. Descriptive statistics about R&D inputs and outputs for 30 OECD countries

Variables	Range	Minimum	Sum	Maximum	Mean	Std. Deviation
Inputs						
RDP(2000) ^a	1288136	1646	3386018	1289782	112867.300	254205.886
RDE(2000) ^b	267548.2	219.25	607031.868	267767.5	20234.396	50872.376
Outputs						
PAPER(2002)	380883	158	1050372	381041	35012.400	70673.129
PATENT(2002)	16017	3	47467	16020	1582.167	3836.958

^a It is a FTE (Full Time Equivalent) on R&D.

^b Its unit is million constant dollars – 2000 prices and PPPs.

This paper uses R&D inputs and outputs with a two-year time lag for 30 OECD countries to serve the relevant quantitative analyses about presented WADD models[®]. The result of

[®] This paper also quotes the empirical data about water supply services in Aida et al. (1998) for a further complementary examination in some necessary situations.

descriptive statistics is displayed in table 1. Evidently, the multivariable evaluation system is characterized by “multi-unit” and “disparate orders of magnitude”.

3. Review on standardization-based formulations of existing WADD models

The basic ADD model was proposed by Charnes et al. (1985), whose specific program formulation (P1) is as follows.

$$\begin{aligned}
& \max \quad \theta = \sum_{i=1}^m s_{io}^- + \sum_{r=1}^s s_{ro}^+ \\
& \text{subject to} \\
& x_{io} = \sum_{j=1}^n \lambda_j x_{ij} + s_i^-, \quad i = 1, 2, \dots, m, \\
& y_{ro} = \sum_{j=1}^n \lambda_j y_{rj} - s_r^+, \quad r = 1, 2, \dots, s, \\
& \sum_{j=1}^n \lambda_j = 1, \\
& 0 \leq \lambda_j, s_i^-, s_r^+.
\end{aligned} \tag{P1}$$

The ADD model is a non-radial and Pareto – Koopmans (PK) model (Cook and Seiford, 2009). The first two equation constraints in program (P1) are translation invariant^④ with the help of the convexity constraint of $\sum_{j=1}^n \lambda_j = 1$ (Cooper et al., 1999), because, for given any real number d_i and c_r , formulas (1) and (2) always hold.

$$s_{io}^- = x_{io} - \sum_{j=1}^n \lambda_j x_{ij} = (x_{io} + d_i) - \sum_{j=1}^n \lambda_j (x_{ij} + d_i), \quad i = 1, 2, \dots, m, \tag{1}$$

$$s_{ro}^+ = \sum_{j=1}^n \lambda_j y_{rj} - y_{ro} = \sum_{j=1}^n \lambda_j (y_{rj} + c_r) - (y_{ro} + c_r), \quad r = 1, 2, \dots, s. \tag{2}$$

Consequently, negative and zero inputs or outputs can be dealt with by adding suitable constants to the affected input or output sets, with assurance that the solutions will not be affected by these translations (Cooper et al., 1999). So, the whole ADD model satisfies translation invariance. However, the sum of slacks in objective function is in-commensurable measure because of the differences in measurement units and magnitude orders of slacks, i.e., the ADD model is not unit invariant (Lovell and Pastor, 1995).

Following the property of unit invariance of transformation formulas (S1) ~ (S5), it's may be a good solution to use a statistical data-based parameter to scale each slack which makes the sum of scaled slacks satisfy the unit invariance. Because these parameters come from practical observation data, they can be seen as objective weights embodying the

^④ Sueyoshi and Sekitani (2009) considered that ADD is not translation invariant considering the occurrence of multiple projections.

status/importance of slacks, which in fact serve as standardization parameters embodying information of corresponding observation data. Thus it's nature to introduce the “weighted” ADD model (WADD), whose objective function is (O1).

$$\theta = \max \sum_{i=1}^m w_i^- s_{io}^- + \sum_{r=1}^s w_r^+ s_{ro}^+ . \quad (O1)$$

In most practical evaluations, w_i^- and w_r^+ are user-specified weights, which embody managers' preference, and are obtained through value judgment (e.g., Thrall, 1996). However, it's difficult to subjectively judge which one of observation indicators is more important than other for the tested DMU because of the supplemental relations between them. Cooper et al. (1999) points out:

“... , it is desirable to orient these choices to more “objective” criteria so that investigators doing scientific research, for instance, will reach the same (or similar) conclusions from the same (or similar) bodies of data. For this purpose, we shall confine ourselves to “data-based weights.”

The data-based weights can embody information hiding in the difference between datasets, and the inefficiency results of WADD models with data-based weights are justifiable and objective. So WADD models with data-based weights on slacks are not only mathematically scientific but also practically applicable.

Of course, for obtaining unit invariance, the weights should be “contragredient” with weighted slacks, which means that the dimensions of weights are reciprocal of these of weighted slacks (Cooper et al., 1999). From economic meaning, weights represent the marginal worth of corresponding weighted slack. Specifically in terms of WADD, weights are associated with unit costs and unit prices of excess and shortfall slack variables (Bardhan et al., 1996), and the sum of weighted slacks is used to measure the total cost of inefficiencies. The modification by weight is to make WADD models satisfy some attractive properties by scaling slacks in order to obtain commensurable inefficiency scores.

3.1 Measure of inefficiency proportions (MIP-WADD)

When $w_i^- = (x_{io})^{-1}$, $w_r^+ = (y_{ro})^{-1}$, the objective function of corresponding WADD model is

$$\theta = \max \sum_{i=1}^m \frac{s_{io}^-}{x_{io}} + \sum_{r=1}^s \frac{s_{ro}^+}{y_{ro}} . \quad (O2)$$

The objective function (O2) is a proportional measure of inefficiency, the WADD with

which is referred to as **MIP-WADD** (Cooper et al., 1999). Because the numerators and denominators in (O2) have same unit, the (O2) satisfies unit invariance. However, MIP-WADD doesn't satisfy translation invariance and suffers from zero and negative data as traditional ADD.

The unequal relationship $s_{io}^- < x_{io}$ for $\forall i$ must hold, but the unequal relationship $s_{ro}^+ < y_{ro}$ for $\forall r$ aren't always true. So, the upper bound of inefficiency score obtained from MIP-WADD possibly exceeds unity, and, similarly, the average^⑤ inefficiency score can't be limited in [0, 1], i.e., can't satisfy the sense of common (in)efficiency.

Besides, in (O2), weights $w_i^- = (x_{io})^{-1}$ and $w_r^+ = (y_{ro})^{-1}$ vary with the tested DMU₀, which can be referred to as "weight instability", so that the ranking of inefficiency measure generated by MIP-WADD may be considered not comparable due to the absence of a common weight for specific observation indicator. In fact, there are differences in unit costs and unit prices of surplus and shortfall slacks for different DMUs, so the unstable weights are, to a certain degree, reasonable in many practical environments.

3.2 WADD with sum-based weight (SUM-WADD)

Ali et al. (1995) also tried to introduce new weights $w_i^- = (\sum_{j=1}^n x_{ij})^{-1}$, $w_r^+ = (\sum_{j=1}^n y_{ij})^{-1}$, where the corresponding objective function is

$$\theta = \max \sum_{i=1}^m \frac{s_{io}^-}{\sum_{j=1}^n x_{ij}} + \sum_{r=1}^s \frac{s_{ro}^+}{\sum_{j=1}^n y_{ij}}. \quad (O3)$$

The WADD model with (O3) is referred to as **SUM-WADD**. The objective function (O3) is the same to (O2) in such key properties as unit and translation invariance. Although it suffers from zero data less possibly than (O2), it can't satisfy some extreme situations, where $\sum_{j=1}^n x_{ij} = 0$ or $\sum_{j=1}^n y_{ij} = 0$ may be exist. However, compared with (O2), the objective function (O3) provides a common stable weight, and makes inefficiency score reduced into [0, 1].

In the light of the estimated density distribution of inefficiency scores (excluding zero scores) of WADD models with (O2) and (O3) by kernel density estimation (Wand and Jones,

^⑤ Usually, $\left[\sum_{i=1}^m (s_{io}^- / x_{io}) + \sum_{r=1}^s (s_{ro}^+ / y_{ro}) \right] / (m + s)$ (Bardhan et al., 1996; Cooper et al., 1999) or $1/m \cdot \sum_{i=1}^m (s_{io}^- / x_{io}) + 1/s \cdot \sum_{r=1}^s (s_{ro}^+ / y_{ro})$ (Cooper et al., 1999) is used as a measure of the average inefficiency.

1995; Simonoff, 1996) (Fig. 1 and Fig. 2) [®], it's easy to find that the variance between two distribution illustrations of inefficiency scores is clear. The inefficiency scores from SUM-WADD mainly converge into (0, 0.02), but majority inefficiency scores from MIP-WADD rang in (0, 10). Clearly, the congestion of inefficiency scores is very serious for SUM-WADD. A common shortcoming of both MIP-WADD and SUM-WADD is that their density function figures are very sharp in a relative smaller range of inefficiency scores against a whole range of inefficiency scores, meaning crowdedness of majority inefficiency scores and shortness of discrimination power.

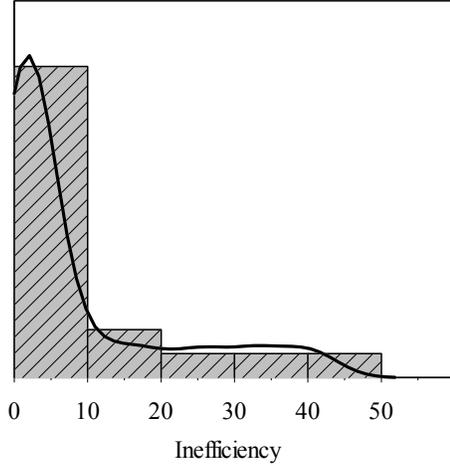


Fig. 1. Kernel density estimation of inefficiency for MIP-WADD

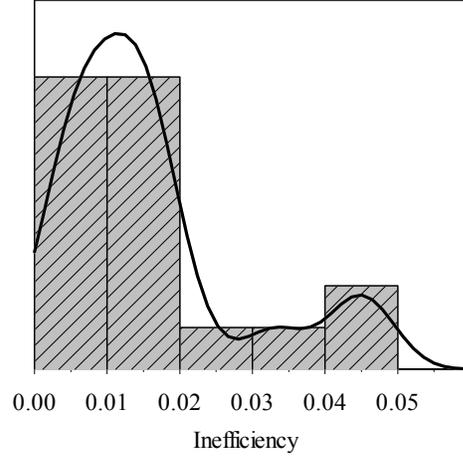


Fig. 2. Kernel density estimation of inefficiency for SUM-WADD

3.3 Range adjusted measure (RAM-WADD)

Among the exiting typical WADD models, NOR-WADD and RAM-WADD can satisfy unit and translation invariance simultaneously.

In a working paper (Cooper and pastor, 1995), **RAM-WADD** is proposed, which can be also found in Cooper et al. (1999).

Here, $w_i^- = (x_i^U - x_i^L)^{-1}$, $w_r^+ = (y_r^U - y_r^L)^{-1}$, so the objective function of RAM-WADD is

$$\theta = \max \left(\sum_{i=1}^m \frac{s_{io}^-}{x_i^U - x_i^L} + \sum_{r=1}^s \frac{s_{ro}^+}{y_r^U - y_r^L} \right). \quad (O4)$$

Because,

$$(x_i^U - d_i) - (x_i^L - d_i) = x_i^U - x_i^L \quad \text{and} \quad (y_r^U - c_r) - (y_r^L - c_r) = y_r^U - y_r^L,$$

hold for $\forall i, r, j$, (O4) is translation invariant (Cooper et al., 1999). However, as shown in the

[®] All distribution figures of inefficiency results in study are estimated by Splus 9.0.

Fig. 3, the inefficiency scores obtained from RAM-WADD mainly converge into (0, 0.04). This is mainly due to big differences between maximum and minimum within the dataset, which causes weights too small. Inefficiency congestion in SUM-WADD is attributable to the same reason.

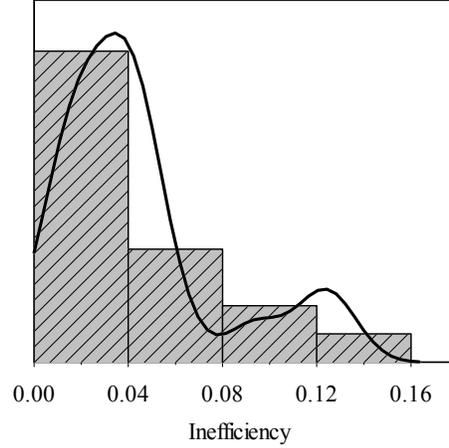


Fig. 3. Kernel density estimation of inefficiency for RAM-WADD

3.4 Standardized additive DEA (NOR-WADD)

Another inefficiency measure which satisfies translation invariance is **NOR-WADD**, which was proposed by Lovell and Pastor (1995).

Here, $w_i^- = (\sigma_i^-)^{-1}$, $w_r^+ = (\sigma_r^+)^{-1}$, and the objective function is

$$\theta = \max \left(\sum_{i=1}^m \frac{s_{io}^-}{\sigma_i^-} + \sum_{r=1}^s \frac{s_{ro}^+}{\sigma_r^+} \right). \quad (O5)$$

Because,

$$\sqrt{\sum_{j=1}^n \left((x_{ij} - d_i) - (\bar{x}_i - d_i) \right)^2 / n - 1} = \sqrt{\sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 / n - 1} = \sigma_i^- \quad \text{and}$$

$$\sqrt{\sum_{j=1}^n \left((y_{ij} - c_r) - (\bar{y}_r - c_r) \right)^2 / n - 1} = \sqrt{\sum_{j=1}^n (y_{ij} - \bar{y}_r)^2 / n - 1} = \sigma_r^+$$

hold for any given constant d_i and c_r , $\forall i, r, j$, (O5) is translation invariant. Comparing density distribution figures of inefficiency scores by NOR-WADD and RAM-WADD (Fig. 3 and Fig. 4), their figures of density function are almost same. The distribution range of inefficiency scores generated by NOR-WADD obviously are larger than RAM-WADD, but majority of inefficiency scores still are belonged to a relative small range (0, 0.2), which usually brings a poor power in completely discriminating inefficiency scores.

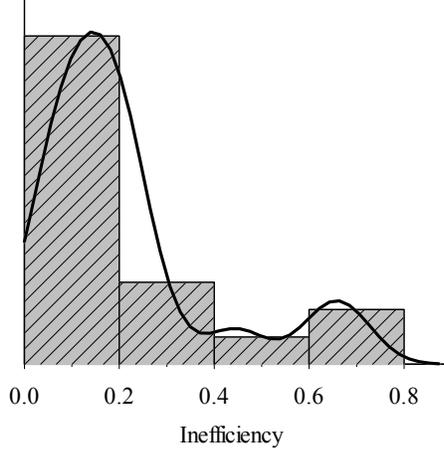


Fig. 4. Kernel density estimation of inefficiency for NOR-WADD

3.5 Rang of improvement measure (RIM-WADD) [Ⓣ]

Besides, Pastor (1994) proposed another inefficiency measure model from the range of possible improvement in each observation indicator (input or output) for DMU₀ (Pastor and Ruiz, 2007) (referred to as **RIM-WADD**).

Because $s_{io}^- = x_{io} - \sum_{j=1}^n \lambda_j x_{ij} \leq x_{io} - x_i^L$ and $s_{ro}^+ = \sum_{j=1}^n \lambda_j y_{rj} - y_{ro} \leq y_r^U - y_{ro}$ hold for $\forall i, r, j$, if let x_{io}^* and y_{ro}^* be contributing inputs and obtained outputs in practical production process, then $\frac{s_{io}^-}{x_{io} - x_i^L} = \frac{x_{io} - x_{io}^*}{x_{io} - x_i^L}$ show the rate of the rang of practical surplus $x_{io} - x_{io}^*$ (potential improvement, i.e., inputs contraction) to the range of possible maximum surplus $x_{io} - x_i^L$, and $\frac{s_{ro}^+}{y_r^U - y_{ro}} = \frac{y_{ro}^* - y_{ro}}{y_r^U - y_{ro}}$ show the rate of the rang of practical shortfall $y_{ro}^* - y_{ro}$ (potential improvement, i.e., output expansion) to the range of possible maximum shortfall $y_r^U - y_{ro}$.

So for RIM-WADD, $w_i^- = (x_{io} - x_i^L)^{-1}$, $w_r^+ = (y_r^U - y_{ro})^{-1}$, i.e., the objective function is

$$\theta = \max \left(\sum_{i=1}^m \frac{s_{io}^-}{x_{io} - x_i^L} + \sum_{r=1}^s \frac{s_{ro}^+}{y_r^U - y_{ro}} \right). \quad (O6)$$

Because $(x_{io} - d_i) - (x_i^L - d_i) = x_{io} - x_i^L$ and $(y_r^U - c_r) - (y_{ro} - c_r) = y_r^U - y_{ro}$ hold for $\forall i, r, j$, where d_i and c_r are any given constant numbers, (O6) also satisfies translation invariance. But, (O6) easily suffers from zero data, because $x_{io} = x_i^L$ and $y_{ro} = y_r^U$ always

[Ⓣ] This model is also extended to accept any returns to scale by Cooper et al. (2011), which is called as ‘‘BAM’’.

exist.

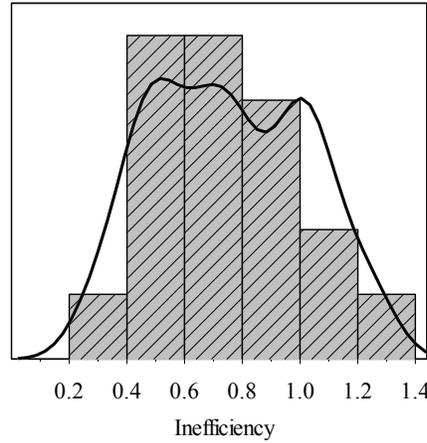


Fig. 5. Kernel density estimation of inefficiency for RIM-WADD

Another problem associated with RIM-WADD is that it is an inefficiency measurement with unstable weights, which is the same to MIP-WADD. However, inefficiency scores generated by it most lie in a larger bounded range (0.4, 1.0) (see Fig. 5), which shows a better power in inefficiency discrimination than SUM-WADD, NOR-WADD and RAM-WADD.

3.6 Brief summary and discussion

Firstly if visually compare the formulas for standardization with the formulations of objective functions of WADD, it's evident that there are corresponding relationships mainly in denominators, such between (S1) and (O5), between (S3) and (O4) as well as between (S5) and (O3). Although, there are no corresponding standardization formulas for (O2) and (O6), we can use their denominators to scale observation datasets for convenient calculation.

Specifically comparing denominators of objective functions (O2) ~ (O3) with (O4) ~ (O6), clearly their formulations are different. The former is a single statistic (including specific observation data, hereafter), however the latter is a difference between two statistics which makes denominators are stable when two statistics add or reduce a constant simultaneously. If let $S(v_i)^a$ and $S(v_i)^b$ standard for the two different statistics (maximum, minimum or mean) of all observation values in the dataset v_i , when adding a constant d_i to the dataset v_i , then

$$S(v_i + d_i)^a - S(v_i + d_i)^b = (S(v_i)^a + d_i) - (S(v_i)^b + d_i) = S(v_i)^a - S(v_i)^b.$$

So the difference between two statistics can make the objective functions satisfy translation invariance, however a single statistic can't. Note that the statistic “range” of

observation dataset is the difference form in itself.

Moreover, considering the constraints in (P1), then

$$s_{io}^- = x_{io} - \sum_{j=1}^n \lambda_j x_{ij} \leq x_{io} - x_i^L \leq x_i^U - x_i^L \Leftrightarrow s_{io}^- / (x_{io} - x_i^L) \leq 1 \text{ and } s_{io}^- / (x_i^U - x_i^L) \leq 1,$$

$$s_{ro}^+ = \sum_{j=1}^n \lambda_j y_{rj} - y_{ro} \leq y_r^U - y_{ro} \leq y_r^U - y_r^L \Leftrightarrow s_{ro}^+ / (y_r^U - y_{ro}) \leq 1 \text{ and } s_{ro}^+ / (y_r^U - y_r^L) \leq 1.$$

The formations of (O4) and (O6) mainly depend on these equivalences, clearly which ensure the objective function values of WADD models range in $[0, 1]$ with a multiplier $1/m + s$. But WADD models with (O2) and (O5) can't ensure that (in)efficiency scores are below unity, so (O4) and (O6) have more superiority than them in satisfying the sense of (in)efficiency.

Besides, for (O3), (O4) and (O5), w_i^- and w_r^+ don't vary from one DMU to another. However, for (O2) and (O6), w_i^- and w_r^+ vary from one DMU to another because of the impact of specific observation. This means that the weights in (O2) and (O6) are unstable. Cooper et al. (1999) meant that (in)efficiency results based on unstable weights can't be used as ranking comparison. Even in the case of (O5), it's not always suitable for all evaluation principles, because the numerators and denominators would be measuring distances from different points—with the standard deviations measuring distance from the mean, while the slack values would be measuring distance from a point on the boundary of the observations (Cooper et al., 1999).

Finally, on formulating WADD models, it's more advisable to use the reciprocal of descriptive statistics as weights in order to obtain unit invariance and even translation invariance, such as NOR-WADD, RAM-WADD and RIM-WADD. In fact, these weights show marginal worth of each unit with nonzero input surplus and output shortfall, however, it is difficult to interpret and justify which descriptive statistics are more suitable to be used as the reciprocal of specific weights. For example, Ali et al. (1995) canceled the mean to scale slacks, and Cooper et al. (1999) thought it not proper to scale slacks using standard deviation in NOR-WADD (Lovell and pastor, 1995) and using the difference between observation data and maximum or minimum in RIM-WADD (Pastor, 1994). The reason of divergent opinions is that they made choices of weights from different perspectives. However, as in this study, it may be more feasible and advisable to implement the interpretation with regard to the standardization of original datasets. This means that it's rational to understand WADD models from the pre-processing of original datasets. This is one main purpose in this study.

4. New extensions of WADD models based on two standardization techniques

4.1 WADD with maximum as weight (Max-WADD)

In many practical statistical analyses, the maximum in a observation dataset usually is used to scale all observation data, i.e., normalizing the original dataset by (S2), in order to make all observation datasets reduced into a commensurable range $[0, 1]$ when original observation data are not negative. So, the objective function of corresponding Max-WADD model can be formulated as (O7) as follows.

$$\theta = \max \left(\sum_{i=1}^m \frac{s_{io}^-}{x_i^U} + \sum_{r=1}^s \frac{s_{ro}^+}{y_r^U} \right). \quad (O7)$$

Evidently, $w_i^- = (x_i^U)^{-1}$ and $w_r^+ = (y_r^U)^{-1}$. The attractive properties of (O7) can be interpreted from two aspects. One aspect is that $s_{io}^-/x_i^U \leq 1$ and $s_{ro}^+/y_r^U \leq 1$ can make θ bounded not exceed unity with a multiplier $1/m+s$, and the other is that $x_i^U \leq \sum_{j=1}^n x_{ij}$ and $y_r^U \leq \sum_{j=1}^n y_{ij}$ make (O7) generate larger proportions than (O3) at least.

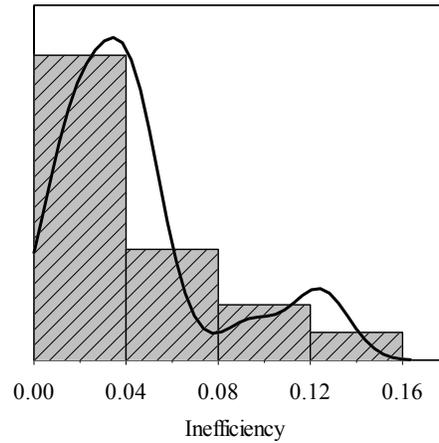


Fig. 6. Kernel density estimation of inefficiency for Max-WADD

Although $x_i^U - x_i^L \leq x_i^U$ and $y_r^U - y_r^L \leq y_r^U$, the inefficiency distribution in Fig. 6 shows that the inefficiency result of Max-WADD is almost the same to that of RAM-WADD. The coincidence also can be proved by the empirical datasets in Aida et al. (1998), which is mainly due to the large difference between maximum and minimum.

4.2 WADD with mean as weight (Mean-WADD)

Based on the analyses about WADD models above, it's clear that inefficiency crowdedness is due to small weights. Especially in many practical applications, because of

different orders of magnitude of datasets (x_{ij} and y_{rj}) for different production scales (as the results in Aida et al. (1998)), there is a big difference between some statistics such as sum, maximum as well as range and majority observation data within the dataset, so weights coming from reciprocals of sum, maximum and range must be smaller. This causes a smaller (in)efficiency score. So for obtaining a better power in (in)efficiency discrimination in terms of frequently-used statistics available, we can let $w_i^- = (\bar{x}_i)^{-1}$ and $w_r^+ = (\bar{y}_r)^{-1}$, then the corresponding objective function is

$$\theta = \max \left(\sum_{i=1}^m \frac{S_{io}^-}{\bar{x}_i} + \sum_{r=1}^s \frac{S_{ro}^+}{\bar{y}_r} \right). \quad (O8)$$

Fig. 7 depicts the density distribution of inefficiency scores based on WADD with (O8). Although it is the similar distribution shape to Fig. 3 depicting inefficiency results from SUM-WADD, the range of inefficiency scores expands much (from (0, 0.05) to (0, 1.6)), showing a better distinction power in inefficiency results.

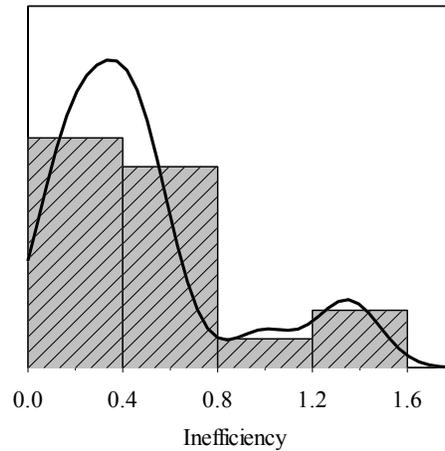


Fig. 7. Kernel density estimation of inefficiency for Mean-WADD

5. An empirical comparison of (in)efficiency discrimination power

By Splus 9.0, the comparison result[®] of Kernel density estimations of inefficiency scores for WADD models with (O3) ~ (O8) gives a complete comparison in inefficiency discrimination power, which is SUM-WADD < RAM (Max)-WADD < NOR-WADD < Mean-WADD < RIM-WADD as shown in Fig. 8.

[®] MIP-WADD is excluded in comparison due to much bigger inefficiency results, otherwise the figures of inefficiency scores based on Kernel density estimations can't provide a clear comparison.

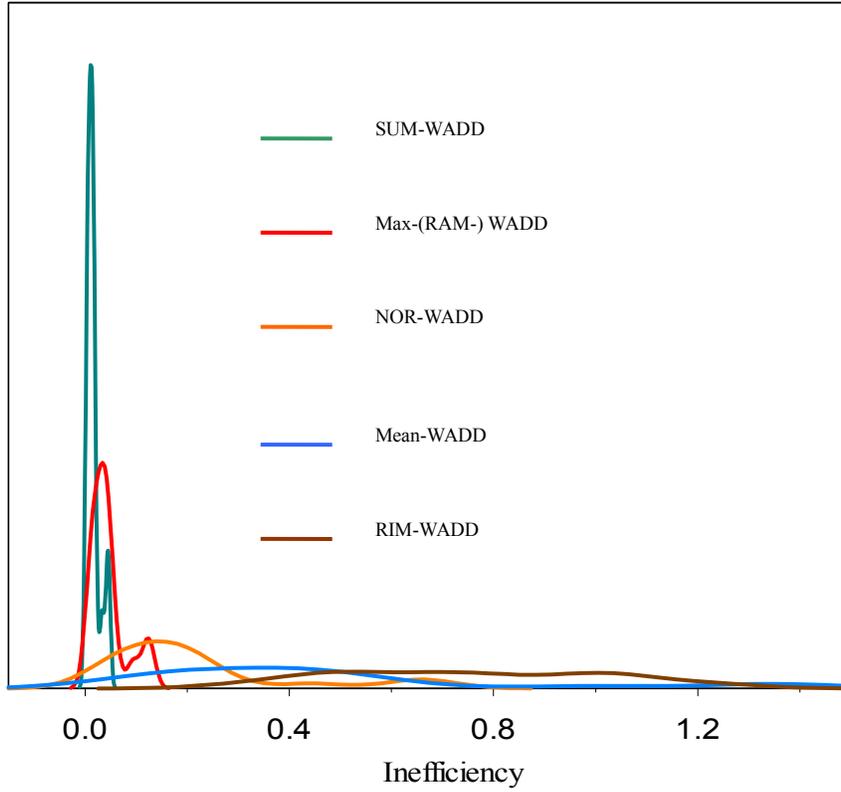


Fig. 8. Comparing Kernel density estimations of inefficiency for (O3) ~ (O7)

Note: The efficiency is calculated by authors associated with Excel solver

Besides, in order to maintain the inefficiency level “ θ^* ” being on the unit scale $[0, 1]$, a constant δ is introduced to balance $\sum_{i=1}^m w_i^- s_{io}^- + \sum_{r=1}^s w_r^+ s_{ro}^+$, i.e.,

$$\theta = \max \delta \left(\sum_{i=1}^m w_i^- s_{io}^- + \sum_{r=1}^s w_r^+ s_{ro}^+ \right).$$

So, if $0 \leq w_i^- s_{io}^- \leq 1$ and $0 \leq w_r^+ s_{ro}^+ \leq 1$, and let $\delta = 1/(m+s)$, then $\theta^* \in [0,1]$ as a measure of inefficiency satisfying the sense of inefficiency. At the same time, a measure of efficiency, $\varphi^* = 1 - \theta^* \in [0,1]$, is obtained, which satisfies the sense of efficiency. Obviously, RAM-WADD, Max-WADD and SUM-WADD can satisfy the condition. In fact, Mean-WADD and NOR-WADD with $\delta = 1/(m+s)$ can satisfy the condition under general environments. The Fig. 9 displays a complete comparison of Kernel density estimations of efficiency scores for WADD models with (O5) ~ (O8) including BCC and SBM (Tone, 2001). Obviously, in discrimination power, BCC (SBM) > RIM-WADD > Mean-WADD > NOR-WADD > RAM/Max -WADD. The efficiency range of BCC and SBM not only is wider, but also distribution curve is not sharp, meaning that efficiency distribution is not centralized

and has a better discrimination power.

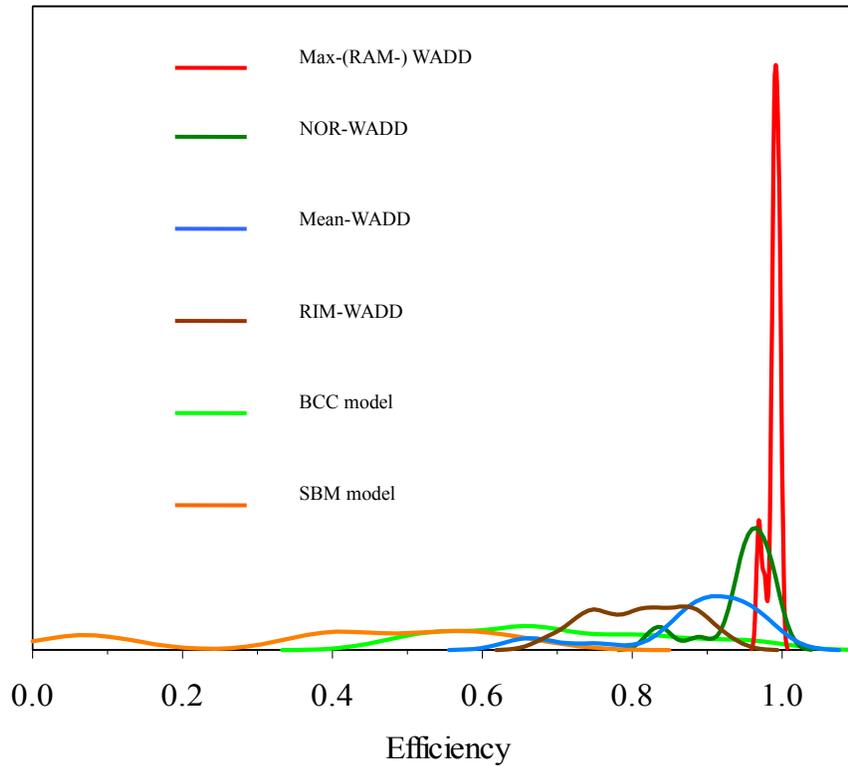


Fig. 9. Comparing Kernel density estimations of efficiencies for (O5) ~ (O8) including BCC and SBM

Note: The efficiency is calculated by authors associated with Excel solver except for BCC efficiency

Fig. 10 provides another illustration for efficiency distribution and discrimination power of WADD models with (O3) ~ (O8) including BCC and SBM by a boxplot, which displays a visual comparison of the variation of efficiency scores across those models. Evidently, the variation of efficiency scores yielded by all WADD models is less than that from BCC and SBM, and Mean-WADD yields a bigger variation than SUM-WADD, Max-WADD, RAM-WADD and NOR-WAD. Among WADD models, the RIM-WADD yields the lowest average efficiency, and there is no visual difference in the average efficiency from SUM-WADD, Max-WADD and RAM-WADD.

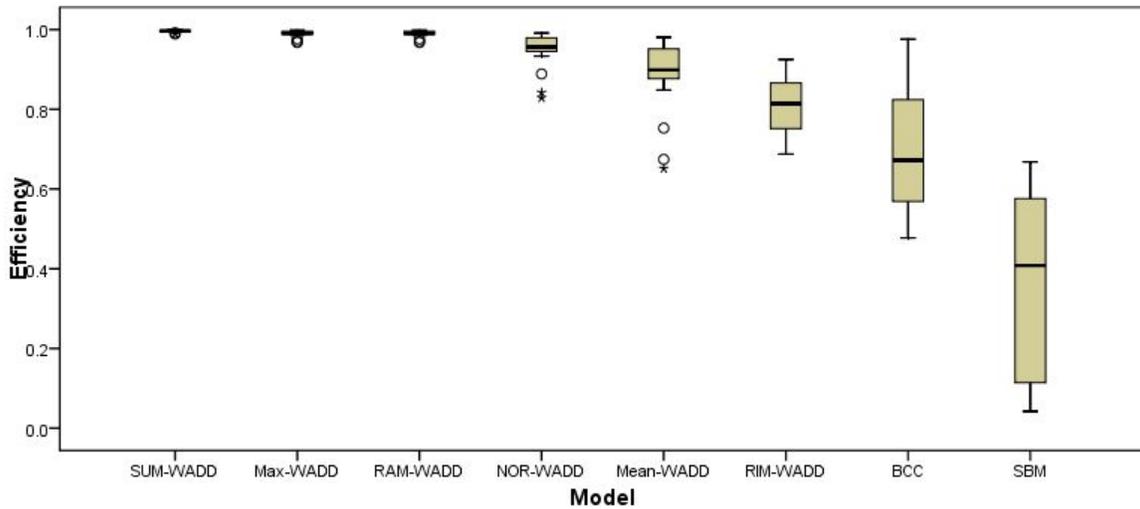


Fig. 10. Comparing efficiency estimates for (O3) ~ (O8) including BCC and SBM

Note: The median for each dataset is indicated by the black centre line (i.e. where 50% of the data is above it and 50% below it). The first and third quartiles of the data are at the edges of the grey box where, respectively, 25% of the data falls below it and 25% of the data is above it. The total length of the grey box is known as the inter-quartile range (IQR). The extreme values (within 1.5 times the inter-quartile range from the upper or lower quartile) are the ends of the lines (or whiskers) extending beyond the grey area. Points that lie at a greater distance from the median than 1.5 times the IQR are plotted individually and may be regarded as outliers.

There is tradeoff between the sense of (in) efficiency with respect to making θ^* satisfying $[0, 1]$ and the discrimination power. Obviously, the bigger values of w_i^- and w_r^+ are needed. However, this always causes inefficiency crowded, which means that some measure formulations are mathematically acceptable but practically unacceptable (Sueyoshi and Sekitani, 2009).

6. Applications

Since this study has related the formulation of WADD models to the standardization of original datasets, it's helpful for users to flexibly apply various WADD models to heterogeneous original datasets and facilitate users' calculations. Furthermore, the attractive translation property of some WADD models helps in improving the formulation of DEA-based hybrid measure model in the presence of zero and negative values.

6.1 Smoothing WADD models' calculation

If w_i^- and w_r^+ come from the statistics of observation datasets, when observation datasets are given, the complete formulation of WADD model for inefficiency measure is as follows.

$$\begin{aligned}
\theta &= \max \sum_{i=1}^m w_i^- s_{io}^- + \sum_{r=1}^s w_r^+ s_{ro}^+ \\
&\text{subject to} \\
x_{io} &= \sum_{j=1}^n \lambda_j x_{ij} + s_i^-, \quad i=1,2,\dots,m, \\
y_{ro} &= \sum_{j=1}^n \lambda_j y_{rj} - s_r^+, \quad r=1,2,\dots,s, \\
\sum_{j=1}^n \lambda_j &= 1, \\
0 &\leq \lambda_j, s_{io}^-, s_{ro}^+.
\end{aligned} \tag{P2}$$

Because w_i^- and w_r^+ are non-zero constants for given sample observation datasets, then

$$\begin{aligned}
x_{io} &= \sum_{j=1}^n \lambda_j x_{ij} + s_{io}^- \Leftrightarrow w_i^- x_{io} = \sum_{j=1}^n \lambda_j w_i^- x_{ij} + w_i^- s_{io}^-, \quad i=1,2,\dots,m, \\
y_{ro} &= \sum_{j=1}^n \lambda_j y_{rj} - s_{ro}^+ \Leftrightarrow w_r^+ y_{ro} = \sum_{j=1}^n \lambda_j w_r^+ y_{rj} - w_r^+ s_{ro}^+, \quad r=1,2,\dots,s.
\end{aligned}$$

If let $\hat{s}_{io}^- = w_i^- s_{io}^- \geq 0$, $\hat{s}_{ro}^+ = w_r^+ s_{ro}^+ \geq 0$, $\hat{x}_{ij} = w_i^- x_{ij}$ and $\hat{y}_{rj} = w_r^+ y_{rj}$, then program (P2) can be changed into

$$\begin{aligned}
\hat{\theta} &= \max \left(\sum_{i=1}^m \hat{s}_{io}^- + \sum_{r=1}^s \hat{s}_{ro}^+ \right) \\
&\text{subject to} \\
\hat{x}_{io} &= \sum_{j=1}^n \lambda_j \hat{x}_{ij} + \hat{s}_{io}^-, \quad i=1,2,\dots,m, \\
\hat{y}_{ro} &= \sum_{j=1}^n \lambda_j \hat{y}_{rj} - \hat{s}_{ro}^+, \quad r=1,2,\dots,s, \\
\sum_{j=1}^n \lambda_j &= 1, \\
0 &\leq \lambda_j, \hat{s}_{io}^-, \hat{s}_{ro}^+.
\end{aligned} \tag{P3}$$

Comparing the original ADD model (P1) and the transformed ADD model (P3), their structures are obviously same. However, program (P3) is dimensionless in both observation data and slacks, so the value of sum formulation $\sum_{i=1}^m \hat{s}_{io}^{-*} + \sum_{r=1}^s \hat{s}_{ro}^{+*}$ is commensurable, which can be used to rank inefficiency level.

In fact, the transformation process from (P1) to (P3) is a standardization (or referred to as scaling) process of original datasets. So, it's necessary to preprocess original datasets when using ADD model.

Remark 1: Before using WADD, one can change heterogeneous original datasets into homogeneous commensurable datasets by corresponding standardization formulas, i.e., scaling input (output) values using corresponding descriptive statistics, and then directly uses

(P3) instead of (P2) to deal with preprocessed data for the comparable efficiency results.

Based on the formulas (1) and (2), for input and output constraints, the formula set (3) holds.

$$\begin{aligned} w_i^- s_{io}^- &= w_i^- x_{io} - \sum_{j=1}^n \lambda_j w_i^- x_{ij} = w_i^- (x_{io} + d_i) - \sum_{j=1}^n \lambda_j w_i^- (x_{ij} + d_i), \quad i = 1, 2, \dots, m. \\ w_r^+ s_{ro}^+ &= \sum_{j=1}^n \lambda_j w_r^+ y_{rj} - w_r^+ y_{ro} = \sum_{j=1}^n \lambda_j w_r^+ (y_{rj} + c_r) - w_r^+ (y_{ro} + c_r), \quad r = 1, 2, \dots, s. \end{aligned} \quad (3)$$

For example, when $w_i^- = (x_i^U - x_i^L)^{-1}$, $w_r^+ = (y_r^U - y_r^L)^{-1}$ and $d_i = -x_i^L$, $c_r = -y_r^L$, the formula set (4) holds for $\forall i, r$.

$$\begin{aligned} \frac{s_{io}^-}{x_i^U - x_i^L} &= \frac{(x_{io} - x_i^L)}{x_i^U - x_i^L} - \sum_{j=1}^n \lambda_j \frac{(x_{ij} - x_i^L)}{x_i^U - x_i^L} \\ \frac{s_{ro}^+}{y_r^U - y_r^L} &= \sum_{j=1}^n \lambda_j \frac{(y_{rj} - y_r^L)}{y_r^U - y_r^L} - \frac{(y_{ro} - y_r^L)}{y_r^U - y_r^L} \end{aligned} \quad (4)$$

When $w_i^- = (\sigma_i^-)^{-1}$, $w_r^+ = (\sigma_r^+)^{-1}$ and $d_i = -\bar{x}_i$, $c_r = -\bar{y}_r$, the formula set (5) holds for $\forall i, r$.

$$\begin{aligned} \frac{s_{io}^-}{\sigma_i^-} &= \frac{(x_{io} - \bar{x}_i)}{\sigma_i^-} - \sum_{j=1}^n \lambda_j \frac{(x_{ij} - \bar{x}_i)}{\sigma_i^-} \\ \frac{s_{ro}^+}{\sigma_r^+} &= \sum_{j=1}^n \lambda_j \frac{(y_{rj} - \bar{y}_r)}{\sigma_r^+} - \frac{(y_{ro} - \bar{y}_r)}{\sigma_r^+} \end{aligned} \quad (5)$$

So, if use ADD model to deal with standardized datasets by (S1) or (S3), the inefficiency results are the same to them from NOR-WADD or RAM-WADD. This is more interpretive and convenient calculated than NOR-WADD or RAM-WADD.

Remark 2: Before using WADD, one can standardize the original observation data directly using (S1) ~ (S5) according to the corresponding formulation of objective function of WADD. The standardized data can be dealt with by ADD which doesn't damage the inefficiency results directly generated by WADD.

This conclusion is supported by Ali et al. (1995). They concluded as follows:

“..... It follows that caution should be exercised in selecting units of measurement if the standard model is to be used without manipulation of the data. To fully ‘standardize’ the model and prevent the inappropriate aggregation of non-commensurable measures, some preprocessing of data may be required.”

6.2 Improving multi-stage models

A multi-stage efficiency measure model was proposed by Fried et al. (1999, 2002) based

on the slacks generated from the BCC-DEA model. It is very attractive and used in many practical fields, such as bank efficiency (Liu and Tone, 2008), R&D efficiency (Wang and Huang, 2007), transit efficiency (Margari et al., 2007) and public institutes efficiency (Glass et al., 2006). The essential contribution of multi-stage model is to generate a “pure” measure for technical efficiency removing the compact of environment factors and stochastic noise, which makes all DMUs compared each other in a common environment and common luck scenario.

However, there are limitations and shortcomings for the multi-stage efficiency measure model. Firstly, in view of the inability dealing with negative and zero values for radial BCC-DEA model due to not being translation invariant, Fried et al. (1999, 2002) has to have proposed adjusted formulations only by adding the difference between maximum predicted slack and predicted slacks on the original inputs and outputs. Besides, BCC slacks adopted in Fried et al. (1999, 2002) are not unit invariant (Lovell and Pastor, 1995) or commensurable. And, BCC model ignores the role of slacks in calibrating the efficiency score, thus measured efficiency allows for upward bias (Liu and Tone, 2008; Cooper et al., 2005). Furthermore BCC slacks are divided into radial and non-radial slacks, which gives rise to concern that some useful information could be lost (Liu and Tone, 2008).

Although Liu and Tone (2008) and Avkiran (2009) proposed new multi-stage measure models based on the SBM model with unit invariance, it still can't deal with negative and zero data due to the shortness of translation invariance in SBM. So, for avoiding the negative and zero values, their adjusted processes still accounted for least favorable operating environment for inputs slacks (most favorable operating environment for output surpluses) and didn't incorporate the most favorable and average operating environments into efficiency measuring.

So an enhanced multi-stage model for making up shortcomings above is necessary. RAM-WADD, RIM-WADD or NOR-WADD has more attracting properties than both BCC model used in Fried et al. (1999, 2002) and SBM model used in Liu and Tone (2008) and Avkiran (2009), which can simultaneously measure the radial and non-radial efficiency and accommodate negative and zero values due to unit and transformation invariance.

Specifically, users firstly apply a chosen WADD model to input and output datasets to obtain an initial efficiency measures of a set of DMUs. The calculation procedure follows the proposed remarks in the above-mentioned section. The improved standardized estimates of

slacks are obtained due to the property of unit-invariance of the chosen WADD model. Then users use stochastic frontier analysis (SFA) to regress those standardized slacks against a set of context variables with consideration of three effects: operating environmental effects, managerial inefficiency, and statistical noise. In the third stage, users adjust standardized inputs or outputs in a manner that accounts for the environmental effects and the statistical noise uncovered in the second stage, and subsequently repeat the first stage analysis by applying the chosen WADD model to the adjusted standardized datasets. A new multistage model for pure efficiency measures, WADD-SFA-WADD, is formed. In our analytical framework, one can adjust original input and output datasets from average, favorable or unfavorable operating environments.

6.3 Brief summary and discussion

It's necessary to standardize original datasets by given transformation formulas in this paper for making them dimensionless and commensurable, or to use corresponding WADD models directly. It's inevitable that there is much heterogeneity in a multivariable evaluation environment, which usually embodies variances in units of measurement across datasets and orders of magnitude within one dataset. So, it's not advisable to aggregate the slacks with heterogeneous measure units directly by sum. Though pre-standardization process by standardization formulas, WADD models with descriptive statistics-based weights can be transformed into the traditional ADD model, the sum of whose slacks can stand for inefficiency level.

Because the constraints of (weighted) additive DEA models meet unit and translation invariance, what one need do is to make objective functions meet the two desirable properties by adding proper weights on slacks. The "statistical data-based" weights should be chosen for a fair and objective comparison. Moreover, the chosen weights should make the value of objective function satisfy a clear discrimination power for comparison requirements, i.e., the formulation of weighted objective function should be mathematically and practically acceptable.

In order to obtain the (in)efficiency results with better discrimination power, the weights should be smaller. Although RAM-WADD and Max-WADD perform well than SUM-WADD, they have poor discrimination power, especially when there are large differences in the production scales of decision-making unites in comparison. When the number of units is large, it's not proper to use the three models to measure (in)efficiency. Our empirical study shows

that Mean-WADD and NOR-WADD usually perform well, and RIM-WADD performs most out of existing WADD models in terms of distribution of (in) efficiency scores.

Of course, the empirical choice of WADD models is not random. Table 2 summarizes the comparisons among the objective functions of seven WADD models in terms of the standardization techniques as well as six properties including (in)efficiency judgment, unit invariance, translation invariance, stability of weight with respect to being invariable or variable for any DMU, sense of (in)efficiency and discrimination power. Table 2 provides us with the fact that none of relevant seven WADD models show a better performance for all six properties. All WADD model perform well in (in)efficiency judgment and unit invariance, but display a inconsistent performance in other four properties. So, decision-makers should choose suitable WADD models depending on the need of specific decision environments.

Table 2. Summary on properties of seven WADD models

Measure model	MIP -WADD	RAM -WADD	NOR -WADD	SUM -WADD	RIM -WADD	Max -WADD	Mean -WADD
Standardization formula-based	No ^a	(S3)	(S1)	(S5)	(S3) or (S3)*	(S2)	(S4)
(In)efficiency judgment	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unit invariance	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Translation invariance	No	Yes	Yes	No	Yes	No	No
Stability of weight	No	Yes	No	Yes	No	Yes	Yes
Sense of (in)efficiency ^b	No	Yes	No	Yes	Yes	Yes	No
Discrimination power ^c	Good	Bad	Relative good	Very bad	Very good	Bad	Good

^a MIP-WADD corresponds to the standardization approach of using the evaluated input and output to scale all inputs and outputs.

^b The sense of (in)efficiency of WADD is obtained associated with scale parameter $\delta = 1/(m + s)$.

^c Discrimination power is based on the distribution of (in)efficiency values associated with Splus 9.0.

Note: (a) “Yes” indicates that a WADD model satisfies a desirable property and “No” indicates an opposite case.

(b) MIP-WADD in Cooper et al. (1999) or Sueyoshi and Sekitani (2009); RAM-WADD in Cooper et al. (1999); NOR-WADD in Lovell and Pastor (1995); SUM-WADD in Ali et al. (1995) or Thrall (1996); RIM-WADD in Pastor (1994); Max-WADD and Mean-WADD in section 4 of this study.

Finally, RAM-WADD, NOR-WADD and RIM-WADD satisfy unit and translation invariance simultaneously, which can deal with negative and zero values. So, they inspire a motivation to improve multi-stage efficiency measure in Fried et al. (1999, 2002) for conquering the negative impact of unit variances of slacks on the efficiency comparison and allowing one to adjust original input and output data from average, favorable or unfavorable operating environments.

7. Conclusion

This paper implements some investigations and comparisons on the formulations and

desirable properties of weighted additive DEA (WADD) associated with the standardization methodologies of original observation datasets. The main purpose of our investigations is to facilitate the applications and extensions of WADD models for users in terms of reasonable interpretations and convenient calculation.

In this study, the formulations of objective functions of weighted additive DEA (WADD) models are endowed with a practical statistical background, which are novelly related to the standardization of original observation datasets, and make WADD models meet the interpretive requirements. It's proved that the preprocessing of original datasets by standardization formulas before using traditional ADD model is the consistent with directly applying the corresponding WADD models. Depending on the translation invariance property of some WADD models, this study propose a new multistage model for pure efficiency measures, WADD-SFA-WADD. By the new hybrid model as an improved analytical framework of Fried et al. (1999, 2002), one can adjust original input and output datasets from average, favorable or unfavorable operating environments.

In terms of future work, our investigations in this study can be applied to various empirical applications. Moreover, it's practically valuable to implement more empirical and theoretical comparisons among WADD models, and investigate more evidences for the choice of WADD models in different evaluation contexts. Especially in the discrimination power of (in)efficiency scores, our rough conclusions in table 2 need more evidences as a future special research task before reaching a general rule.

Reference:

- Aida K, Cooper WW, Pastor JT, Sueyoshi T (1998). Evaluating Water Supply Services in Japan with RAM—A Range-Adjusted Measure of Inefficiency. *Omega* 26:207-232.
- Ali AI, Lerne CS, Seiford LM (1995). Components of efficiency evaluation in data envelopment analysis. *European Journal of Operational Research* 80: 462-473.
- Ali AI, Seiford, LM (1990). Translation invariance in data envelopment analysis. *Operations Research Letters* 9:403-405.
- Avkiran NK (2009). Removing the impact of environment with units-invariant efficient frontier analysis: An illustrative case study with intertemporal panel data. *Omega*, 37:535-544.
- Banker RD (1984) Estimating most productive scale size using data envelopment analysis. *European Journal of Operational Research* 17:35-44.
- Bardhan I, Bowlin WF, Cooper WW, Sueyoshi T (1996). Models and Measures for Efficiency Dominance

- in DEA. Part I: Additive Models and MED Measures. *Journal of the Operations Research Society of Japan* 39:322-332.
- Charnes A, Cooper WW, Golany B, Seiford L, Stutz J (1985). Foundations of Data Envelopment Analysis for Pareto-Koopmans efficient empirical production functions. *Journal of Econometrics* 30:97-107.
- Charnes A, Cooper WW, Rhodes E (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research* 2:429-444.
- Cook WD, Seiford LM (2009). Data envelopment analysis (DEA) – Thirty years on. *European Journal of Operational Research* 192:1-17.
- Cooper WW, Park KS, Pastor JT (1999). RAM: A Range Adjusted Measure of Inefficiency for Use with Additive Models, and Relations to Other Models and Measures in DEA. *Journal of Productivity Analysis* 11:5-42.
- Cooper WW, Pastor JT, Aparicio J, Pastor D (2011). BAM: a bounded adjusted measure of efficiency for use with bounded additive models 2:85-94.
- Cooper WW, Seiford LM, Tone K, Zhu J (2007). Some models and measures for evaluating performances with DEA: past accomplishments and future prospects. *Journal of Productivity Analysis* 28:151-163.
- Fried HO, Lovell CAK, Schmidt SS, Yaisawarng S (2002). Accounting for environmental effects and statistical noise in data envelopment analysis. *Journal of Productivity Analysis* 17:157-174.
- Fried HO, Schmidt, SS, Yaisawarng S (1999). Incorporating the operating environment into a nonparametric measure of technical efficiency. *Journal of Productivity Analysis* 12:249-267.
- Glass JC, McCallion G, McKillop DG, Stringer K (2006). A ‘technically level playing-field’ profit efficiency analysis of enforced competition between publicly funded institutions. *European Economic Review* 50:1601-1626.
- Liu J, Tone K (2008). A multistage method to measure efficiency and its application to Japanese banking industry. *Socio-Economic Planning Sciences* 42:75-91.
- Lovell CAK, Pastor JT (1995). Units invariant and translation invariant DEA models. *Operations Research Letters* 18:147-151.
- Margari BB, Erbetta F, Carmelo P, Piacenza M (2007). Regulatory and environmental effects on public transit efficiency: a mixed DEA-SFA approach. *Journal of Regulatory Economics* 32:131-151.
- Pastor JT (1994). New Additive Models for Handling Zero and Negative Data. Working Paper, Departamento de Estadística e Investigación Operativa, Universidad de Alicante, Spain.
- Pastor JT (1996). Translation invariance in data envelopment analysis: A generalization. *Annals of Operations Research* 66:93-102.
- Pastor JT, Ruiz J (2007). Variables With Negative Values In Dea. In: Zhu J, Cook WD. (Eds.) *Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis*. Springer, New York.
- Simonoff JS (1996). *Smoothing Methods in Statistics*. Springer, New York.
- Sueyoshi T, Sekitani K (2009). An occurrence of multiple projections in DEA-based measurement of technical efficiency: Theoretical comparison among DEA models from desirable properties. *European*

- Journal of Operational Research 196: 764-794.
- Thrall RM (1996). Duality, classification and slacks in DEA. *Annals of Operations Research* 66: 109-138.
- Tone K (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research* 130:498-509.
- Wand MP, Jones MC (1995). *Kernel Smoothing*. Chapman and Hall, London.
- Wang EC, Huang W (2007). Relative efficiency of R&D activities: An across-country study accounting for environmental factors in the DEA approach. *Research Policy* 36:260-273.