CENTRAL LIMIT THEOREMS FOR DIRECTIONAL DISTANCE FUNCTIONS WITH AND WITHOUT UNDESIRABLE OUTPUTS

Léopold Simar, Valentin Zelenyuk, Shirong Zhao

LIDAM Discussion Paper ISBA
2024 / 10

UCLouvain

ISBA | Louvain Institute of Data Analysis and Modeling in economics and statistics

ISBA
Voie du Roman Pays 20 - L1.04.01
B-1348 Louvain-la-Neuve
Email : lidam-library@uclouvain.be
https://uclouvain.be/en/research-institutes/lidam/isba/publication.html

Central Limit Theorems for Directional Distance Functions with and without Undesirable Outputs

Léopold Simar* Valentin Zelenyuk† Shirong Zhao‡

March 4, 2024

Abstract

We develop new central limit theorems (CLTs) for the aggregate directional distance functions (DDFs), which embed the CLTs for the aggregate efficiency and simple mean DDFs as special cases. Moreover, we develop new CLTs for the aggregate DDFs in the presence of the weak disposability of undesirable outputs. Our Monte-Carlo simulations confirm the good performance of statistical inference based on the new CLTs we have derived and illustrate how wrong the inference based on the standard CLTs can be. To our knowledge, this is the first study that provides both the asymptotic theory and the simulation evidence for the non-parametric frontier approaches when some outputs are undesirable. Finally, we provide an empirical illustration using a data set from large US banks as well as supply the computational code for alternative applications.

Keywords: Inference, Data Envelopment Analysis, Non-parametric Efficiency Estimators, Undesirable Outputs, Weak Disposability

JEL Codes: C12, C13, C14

*Simar: Institut de Statistique, Biostatistique et Sciences Actuarielles (ISBA), LIDAM, Université Catholique de Louvain, Voie du Roman Pays 20, B1348 Louvain-la-Neuve, Belgium; email: leopold.simar@uclouvain.be.

†Zelenyuk: School of Economics and Centre for Efficiency and Productivity Analysis (CEPA), University of Queensland, Colin Clark Building (39), St Lucia, Brisbane, Qld 4072, Australia; email: v.zelenyuk@uq.edu.au.

‡Zhao: School of Finance, Dongbei University of Finance and Economics, Dalian, Liaoning 116025, China; email: shironz@163.com.

1 Introduction

Researchers and policymakers often focus on the performance of groups of individuals (e.g., firms), such as the efficiency of firms with different ownerships. This raises the question of estimating the aggregate efficiency of groups of individuals. The topic is inspired by Farrell (1957) and further developed by Färe and Zelenyuk (2003), Simar and Zelenyuk (2007, 2018, 2020), Pham et al. (2024), Simar et al. (2024), to mention a few. The previous literature mainly focused on aggregate efficiency for the Farrell-type efficiency measures, be it in input or output orientation. Recently, Zelenyuk and Panchenko (2023) generalized the theory for the aggregation of directional distance functions (DDFs), which embeds the aggregation of the Farrell-type efficiency metrics as well as the simple mean DDFs as special cases.

The statistical properties for the aggregation of the Farrell-type input and output oriented measures of technical efficiency have also recently been developed by researchers (Simar and Zelenyuk, 2018; Simar et al., 2024), while those for the aggregate DDFs have remained unknown until now. Moreover, to the best of our knowledge, the statistical properties of non-parametric frontier efficiency methods under the weak disposability (i.e. when there are undesirable outputs) have not been explored, neither in theory nor in simulations and so we will try to break the ice on this important aspect of modeling.

In a nutshell, in this paper, we establish the central limit theorems (CLTs) for the aggregation of DDFs with general disposability, including: (i) the strong disposability where all the inputs and outputs are assumed to be freely disposed; (ii) the weak disposability where the undesirable outputs are not freely disposed and its decrease must be accompanied by the decrease of desirable outputs. Our contribution to the literature is twofold: (i) establishing the CLTs for the aggregate DDFs (not only for simple means but also for aggregate means) and (ii) accounting for undesirable outputs. Our statistical results for the aggregation of DDFs further enrich the recent developments of the statistical theories for the efficiency and productivity measures estimated using non-parametric frontier efficiency methods.

Moreover, Monte-Carlo (MC) simulations are employed to assess the performance of the developed CLTs for the aggregate DDFs for different scenarios in the presence of undesirable outputs. These scenarios mimic the true (or hypothetical) technology that involves undesirable outputs in practice. Our MC simulations affirm the effectiveness of the statistical inference built upon the novel CLTs we have formulated. Furthermore, they highlight the potential pitfalls of inference relying on the conventional CLTs (e.g. the Lindeberg-Feller CLT).

We also use the publicly available data set for the large US commercial banks to illustrate the developed theoretical results for the aggregate DDFs. In this illustration, we consider the non-performing loans as the undesirable outputs and model them under weak disposibility. The findings suggest that employing the conventional CLTs for statistical inference in this context could potentially yield misleading results.

The remainder of our paper goes as follows. We lay out the production economics model as well as the definition of the aggregate DDFs in Section 2. Section 3 provides the background for the estimation theory. Section 4 gives an alternative formulation of the directional distance functions and its estimator considering the existence of undesirable outputs, when they are modeled under the weak disposability of desirable and undesirable outputs. Section 5 formally derives the CLTs for the aggregate DDFs. The MC experiments are given in Section 6 to examine the developed theoretical results. Section 7 uses the data of the large US banks to illustrate the theoretical results. Section 8 concludes this paper. Additional results are provided in the online Supplement.

2 The Conceptual Framework

2.1 The Production Model

The typical technology production set can be defined as

$$\mathcal{T} = \{(x, y) \mid \text{outputs } y \in \mathbb{R}^q_+ \text{ can be produced from inputs } x \in \mathbb{R}^p_+\}. \tag{2.1}$$

Equation (2.1) depicts the collection of achievable input-output combinations (x, y) within the pertinent input-output domain. The upper bound of $\mathcal{T}$ is referred to as the *technology frontier*, given by

$$\mathcal{T}^\partial := \{(x, y) \mid \ (x - \delta d_x, y + \delta d_y) \notin \mathcal{T}, \ \forall \, (x, y) \in \mathcal{T}, \delta > 0, d_x \in \mathbb{R}^p_+, d_y \in \mathbb{R}^q_+\}\,. \tag{2.2}$$

The efficiency of a specific DMU at (x, y) can be assessed by calculating the distance between the coordinates in $\mathcal{T}$ and its boundary, $\mathcal{T}^\partial$, in a certain direction $(-d_x, d_y)$, chosen by the researcher. All the points on the technology frontier are deemed as technically efficient and all the points in the interior of $\mathcal{T}$ are not technically efficient, conditional on the direction of measurement $(-d_x, d_y)$. When evaluating the performance of the firms, the common assumptions in the literature (Shephard, 1953, 1970; Färe and Primont, 1995, etc) are used in this paper. Specifically:

Assumption 1. *$\mathcal{T}$ is a closed set.*

Assumption 2. *$\forall\, x \in \mathbb{R}_+^p$, $(x, y) \in \mathcal{T}$, the output set defined as $P(x) = \{y \mid (x, y) \in \mathcal{T}\}$ is bounded.*

Assumption 3. *If $y \geqq 0$, $y \neq 0$, $x = 0$, then $(x, y) \notin \mathcal{T}$; i.e., free lunch is not allowed.*

Assumption 4. *$\forall\, x \in \mathbb{R}_+^p, (x, 0) \in \mathcal{T}$; i.e., zero production is possible.*

Assumption 5. *$\forall\, (x, y) \in \mathcal{T}$, if $y \geqq \ddot{y} \geqq 0$ and $x \leqq \ddot{x}$, then $(\ddot{x}, \ddot{y}) \in \mathcal{T}$; i.e., both outputs and inputs are strongly disposable.*

The efficiency measure based on the DDF for a DMU with an input-output combination (x, y) (Chambers et al., 1996, 1998) is given by

$$\delta(x, y \mid d_x, d_y, \mathcal{T}) := \sup\{\delta \in \mathbb{R}^1 \mid (x - \delta d_x, y + \delta d_y) \in \mathcal{T}\}, \tag{2.3}$$

where $(-d_x, d_y) \in \mathbb{R}_-^p \times \mathbb{R}_+^q$ specifies the projection direction that moves (x, y) to the technology frontier. Note that all the points lying on the frontier have zero values for $\delta(x, y \mid d_x, d_y, \mathcal{T})$ while all the points lying in the interior of $\mathcal{T}$ have positive values for $\delta(x, y \mid d_x, d_y, \mathcal{T})$. There are many possible choices for the direction vector $(-d_x, d_y)$, such as a specific direction for each producer or a common one for all producers. Moreover, we can recover the Farrell-type input and output oriented measures of technical efficiency from $\delta(x, y \mid d_x, d_y, \mathcal{T})$ upon a certain directional vector is specified. For example, $\delta(x, y \mid 0, y, \mathcal{T}) = \lambda(x, y \mid \mathcal{T}) - 1$ and $\delta(x, y \mid x, 0, \mathcal{T}) = 1 - \theta(x, y \mid \mathcal{T})$, where $\lambda(x, y \mid \mathcal{T}) \in [1, \infty)$ and $\theta(x, y \mid \mathcal{T}) \in (0, 1]$ are the Farrell-type output and input efficiency, respectively.[1]

When the firms also produce undesirable/bad outputs, the assumption of free (strong) disposability of all (desirable and undesirable) outputs (i.e., Assumption 5) is not reasonable

[1] See Chambers et al. (1996, 1998) for more details on this measure.

(Färe et al., 1989), whenever it is not free (i.e., incurs a cost) to cut down the undesirable outputs. In order to account for this scenario, let $y = (g, b)$ and $q = q_g + q_b$, where $g \in \mathbb{R}_+^{q_g}$ and $b \in \mathbb{R}_+^{q_b}$ are, respectively, desirable and undesirable outputs. With the existence of undesirable outputs, following the literature (Shephard, 1970; Färe et al., 1989), we replace Assumption 5 by the following two assumptions,

Assumption 6. *$\forall\, (x, g, b) \in \mathcal{T}$, if $g \geqq \ddot{g} \geqq 0$ and $x \leqq \ddot{x}$, then $(\ddot{x}, \ddot{g}, b) \in \mathcal{T}$; i.e., both desirable outputs and inputs are strongly disposable.*

Assumption 7. *$\forall\, (x, g, b) \in \mathcal{T}$, if $\alpha \in [0, 1]$, then $(x, \alpha g, \alpha b) \in \mathcal{T}$; i.e., desirable and undesirable outputs are jointly weak disposable.*

Assumption 7 indicates that to reduce undesirable outputs, it is necessary to decrease desirable outputs in the same (or less desirable) proportion.[2]

The DDF in (2.3) can also be extended to the case with undesirable outputs as

$$\delta(x, g, b \mid d_x, d_g, d_b, \mathcal{T}) := \sup\{\delta \in \mathbb{R}^1 \mid (x - \delta d_x, g + \delta d_g, b - \delta d_b) \in \mathcal{T}\}, \tag{2.4}$$

where the vector $(-d_x, d_g, -d_b) \in \mathbb{R}_-^p \times \mathbb{R}_+^{q_g} \times \mathbb{R}_-^{q_b}$ specifies the projection direction that moves (x, g, b) to the frontier of the technology. Thus, $\delta(x, g, b \mid d_x, d_g, d_b, \mathcal{T})$ measures the maximal value of decrease in both undesirable outputs and inputs as well as simultaneous increase in desirable outputs along the specified directional vector $(-d_x, d_g, -d_b)$.

2.2 Aggregate DDFs

Following Zelenyuk and Panchenko (2023), the aggregation of DDFs for a set of n DMUs can be done via

$$\Delta := \sum_{i=1}^{n} \delta(x_i, y_i \mid d_x^i, d_y^i, \mathcal{T}) S_d^i. \tag{2.5}$$

Here S_d^i represents the weight of the ith DMU, given by

$$S_d^i = \frac{w_y d_y^i + w_x d_x^i}{\sum_{i=1}^{n} (w_y d_y^i + w_x d_x^i)}, \tag{2.6}$$

[2] See Shephard (1970), Färe et al. (1989), and Pham and Zelenyuk (2019) for more discussion.

where $w_x \in \mathbb{R}^p_+$ and $w_y \in \mathbb{R}^q_+$ are the input and output prices (respectively), and $(-d^i_x, d^i_y)$ is the individual specified directional vector. Note that the weight S^i_d is derived from assumptions on the optimization behavior of firms and depends on the input and ouptut prices as well as the chosen individual directional vector. Consequently, the aggregation of DDFs, Δ, also depends on the chosen directional vectors for all the producers.

In the presence of undesirable outputs, we first partition y_i as $y_i = (g_i, b_i) \in \mathbb{R}^{q_g}_+ \times \mathbb{R}^{q_b}_+$, d^i_y as $d^i_y = (d^i_g, -d^i_b) \in \mathbb{R}^{q_g}_+ \times \mathbb{R}^{q_b}_-$, and w_y as $w_y = (w_g, -w_b) \in \mathbb{R}^{q_g}_+ \times \mathbb{R}^{q_b}_-$.[3] We then can extend the definition of the aggregation of DDFs in Zelenyuk and Panchenko (2023) as

$$\Delta = \sum_{i=1}^{n} \delta(x_i, g_i, b_i \mid d^i_x, d^i_g, d^i_b, \mathcal{T}) S^i_d, \tag{2.7}$$

where S^i_d represents the weight of the ith DMU, given by

$$S^i_d = \frac{w_g d^i_g + w_b d^i_b + w_x d^i_x}{\sum_{i=1}^{n}(w_g d^i_g + w_b d^i_b + w_x d^i_x)}. \tag{2.8}$$

To save space and to avoid confusions, in the following equations we will use y_i to represent the outputs and it will be partitioned as desirable and undesirable outputs when needed. Following the discussion in Zelenyuk and Panchenko (2023), next we consider several specific directions for Δ in (2.5) to clarify how it embeds various efficiency measures as special cases.[4]

2.2.1 The Case of Input Orientation: $(d^i_x, d^i_y) = (x_i, 0)$

Given the directional vector $(d^i_x, d^i_y) = (x_i, 0)$, we have $\delta(x_i, y_i \mid d^i_x, d^i_y, \mathcal{T}) = 1 - \theta(x_i, y_i \mid \mathcal{T})$, where $\theta(x_i, y_i \mid \mathcal{T}) \leq 1$ is the Farrell-type input oriented measure of technical efficiency (Farrell, 1957), defined as

$$\theta(x_i, y_i \mid \mathcal{T}) := \inf\{\theta \in \mathbb{R}^1_{++} \mid (\theta x_i, y_i) \in \mathcal{T}\}. \tag{2.9}$$

Thus, equation (2.5) reduces to

$$\Delta = 1 - \sum_{i=1}^{n} \theta(x_i, y_i \mid \mathcal{T}) S^i_d, \tag{2.10}$$

[3] The prices for the undesirable outputs are assumed to be non-positive, reflecting the fact that it is not free to reduce the undesirable outputs. Moreover, for most cases in practice they are not observed and might need to be estimated, see the discussions in Färe et al. (1993), Färe et al. (2005) and Färe et al. (2006).

[4] More details can be found in Table 1 of Zelenyuk and Panchenko (2023), where they list 13 special cases in total.

where

$$S_d^i = \frac{w_x x_i}{\sum_{i=1}^n w_x x_i}. \tag{2.11}$$

Moreover, $\sum_{i=1}^n \theta(x_i, y_i \mid \mathcal{T}) S_d^i$ is the aggregate input efficiency proposed in Färe and Zelenyuk (2003). It is clear that the aggregate DDFs embed the aggregate input efficiency as a special case.[5]

2.2.2 The Case of Output Orientation: $(d_x^i, d_y^i) = (0, y_i)$

Given the directional vector $(d_x^i, d_y^i) = (0, y_i)$, we have $\delta(x_i, y_i \mid d_x^i, d_y^i, \mathcal{T}) = \lambda(x_i, y_i \mid \mathcal{T}) - 1$, where $\lambda(x_i, y_i \mid \mathcal{T}) \geq 1$ is the Farrell-type output oriented measure of technical efficiency (Farrell, 1957), given by

$$\lambda(x_i, y_i \mid \mathcal{T}) := \sup\{\lambda \in \mathbb{R}^1_{++} \mid (x_i, \lambda y_i) \in \mathcal{T}\}. \tag{2.12}$$

Thus, equation (2.5) reduces to

$$\Delta = -1 + \sum_{i=1}^n \lambda(x_i, y_i \mid \mathcal{T}) S_d^i, \tag{2.13}$$

where

$$S_d^i = \frac{w_y y_i}{\sum_{i=1}^n w_y y_i}. \tag{2.14}$$

Moreover, $\sum_{i=1}^n \lambda(x_i, y_i \mid \mathcal{T}) S_d^i$ is the aggregate output efficiency proposed in Färe and Zelenyuk (2003). It is clear that the aggregate DDFs embed the aggregate output efficiency as a special case.[6]

We can also extend (2.5) to the case of the output orientation with desirable and undesirable outputs by considering two cases for the directional vector. For the first case, given the directional vector $(d_x^i, d_g^i, d_b^i) = (0, g_i, b_i)$, equation (2.5) reduces to

$$\Delta = \sum_{i=1}^n \delta(x_i, g_i, b_i \mid 0, g_i, b_i, \mathcal{T}) S_d^i, \tag{2.15}$$

[5] The CLTs for the aggregate input efficiency have been established and also further improved in large dimensions and finite sample sizes by Simar et al. (2024).

[6] The CLTs for the aggregate output efficiency have been established by Simar and Zelenyuk (2018) and further improved by Simar and Zelenyuk (2020), Nguyen et al. (2022), and Simar et al. (2024).

where

$$S_d^i = \frac{w_g g_i + w_b b_i}{\sum_{i=1}^n (w_g g_i + w_b b_i)}. \tag{2.16}$$

Moreover, $\delta(x_i, g_i, b_i \mid 0, g_i, b_i, \mathcal{T})$ is somewhat similar to the Farrell-type output efficiency and the so-called hyperbolic technical efficiency, because it measures the maximal proportional decrease in undesirable outputs as well as the simultaneous increase in desirable outputs by the same proportion, while keeping all inputs and the technology fixed.

It might be also interesting to represent it in terms of the Farrell-type technical efficiency with the undesirable-output orientation, defined analogously to the Farrell-type technical efficiency with input or output orientation, but contracting all undesirable outputs. For the second case, given the directional vector $(d_x^i, d_g^i, d_b^i) = (0, 0, b_i)$, equation (2.5) reduces to

$$\Delta = \sum_{i=1}^n \delta(x_i, g_i, b_i \mid 0, 0, b_i, \mathcal{T}) S_d^i, \tag{2.17}$$

where

$$S_d^i = \frac{w_b b_i}{\sum_{i=1}^n w_b b_i}. \tag{2.18}$$

Moreover, $\delta(x_i, g_i, b_i \mid 0, 0, b_i, \mathcal{T})$ measures the maximal proportional decrease in undesirable outputs given the desirable outputs, inputs and the technology fixed.

2.2.3 The Case of One Specific Orientation: $(d_x^i, d_g^i, d_b^i) = (x_i, 0, b_i)$

Given the directional vector $(d_x^i, d_g^i, d_b^i) = (x_i, 0, b_i)$, equation (2.5) reduces to

$$\Delta = \sum_{i=1}^n \delta(x_i, g_i, b_i \mid x_i, 0, b_i, \mathcal{T}) S_d^i, \tag{2.19}$$

where

$$S_d^i = \frac{w_x x_i + w_b b_i}{\sum_{i=1}^n (w_x x_i + w_b b_i)}, \tag{2.20}$$

and $\delta(x_i, g_i, b_i \mid x_i, 0, b_i, \mathcal{T})$ represents the maximal proportional decrease in inputs and undesirable outputs, while keeping desirable outputs and the technology constant.[7]

[7] The case of this specific orientation was considered in the empirical illustration.

2.2.4 The Case of A Fixed Direction

When the directional vector is fixed for all observations, i.e., $(d_x^i, d_y^i) = (d_x, d_y)$ for all i, which includes the unitary directions, i.e., $(d_x, d_y) = (1, 1)$, the weight reduces to $S_d^i = 1/n$. Equation (2.5) reduces to

$$\Delta = \frac{1}{n} \sum_{i=1}^{n} \delta(x_i, y_i \mid d_x, d_y, \mathcal{T}). \tag{2.21}$$

Thus, here the aggregate DDFs embed the simple mean DDF as a special case.

3 Estimation Theory Background

3.1 Estimators of Individual Efficiency: the Standard Case

For the empirical analysis, however, we do not observe $\delta(x, y \mid d_x, d_y, \mathcal{T})$ and therefore it needs to be estimated using the sample data. The theory and the simulation framework we present in this paper should be valuable (with appropriate adaptations) to a range of approaches, such as Data Envelopment Analysis (DEA), Stochastic Frontier Analysis (SFA), Free Disposal Hull (FDH), or their hybrids.[8] Considering them all in one paper is clearly infeasible, so we will focus on the adaptations to one of them, the DEA approach, which originated in the seminal works of Farrell (1957) and Charnes et al. (1978) and elaborated further and applied in many studies afterwards.[9]

Consider $\mathcal{X}_n = \{(X_i, Y_i)\}_{i=1}^n$ as a sample consisting of the observed inputs and outputs for n observations. If Assumption 5 is imposed, i.e., both outputs and inputs are freely disposable, and if the convexity and CRS (constant returns to scale) property for the production set $\mathcal{T}$ are assumed, the CRS-DEA estimator can be used, which is defined as

$$\widehat{\delta}_{\text{CRS}}(x, y \mid d_x, d_y, \mathcal{X}_n) = \max_{\delta, s_1, \ldots, s_n} \left\{ \delta \geq 0 \mid (y + \delta d_y) \leqq \sum_{i=1}^{n} s_i Y_i,\ (x - \delta d_x) \geqq \sum_{i=1}^{n} s_i X_i,\ \forall\ s_i \geq 0 \right\}. \tag{3.1}$$

[8] For different SFA-type approaches, e.g., see Aigner et al. (1977), Schmidt and Sickles (1984), Greene (2005), Centorrino and Parmeter (2024), and Chapters 11–16 in Sickles and Zelenyuk (2019) for a textbook-style review and many more references therein.

[9] For DEA-type approaches, e.g., see Banker et al. (1984), Deprins et al. (1984), Bogetoft (1994), Olesen and Petersen (1995), Korostelev et al. (1995), Ruggiero (1996), Kuosmanen (2005), Kuosmanen and Podinovski (2009), Podinovski (2022) and a comprehensive review in Ray (2004) and Chapter 8–10 of Sickles and Zelenyuk (2019) and many more references therein.

Alternatively, if $\mathcal{T}$ still exhibits convexity but may first exhibit increasing returns, possibly turning to some local constant returns to scale and eventually to decreasing returns, then one could use the VRS-DEA estimator given by

$$\begin{aligned}\widehat{\delta}_{\mathrm{VRS}}(x, y \mid d_x, d_y, \mathcal{X}_n) = \max_{\delta, s_1, \ldots, s_n} \Big\{ \delta \geq 0 \mid (y + \delta d_y) \leqq \sum_{i=1}^{n} s_i Y_i,\ (x - \delta d_x) \geqq \sum_{i=1}^{n} s_i X_i, \\ \sum_{i=1}^{n} s_i = 1,\ \forall\ s_i \geq 0 \Big\}. \end{aligned} \tag{3.2}$$

The above estimators of $\delta(x, y \mid d_x, d_y, \mathcal{T})$ demonstrate consistency with a convergence rate of n^{κ}, where $\kappa = 2/(p+q)$ if the CRS-DEA estimator is used, or $\kappa = 2/(p+q+1)$ if the VRS-DEA estimator is applied. See Simar et al. (2012), Simar and Vanhems (2012) as well as Simar and Wilson (2015) for a recent review. Moreover, all the estimators are biased with the order as $O\left(n^{-\kappa}\right)$. Finally, Kneip et al. (2015) developed CLTs for the simple mean Farrell-type input efficiency, and this framework can be further generalized to the simple mean DDFs.

3.2 Estimators of Individual Efficiency: the Undesirable Outputs Case

If Assumption 6 and Assumption 7 are imposed instead of Assumption 5 , i.e., apart from assuming strong disposability of inputs and desirable outputs, the desirable and undesirable outputs exhibit joint weak disposability, the standard DEA estimators discussed above do not apply. The DEA-type estimation of $\mathcal{T}$ has been discussed by Färe et al. (1989), Färe et al. (1994), Färe and Grosskopf (2003), Kuosmanen (2005), Färe and Grosskopf (2009), Kuosmanen and Podinovski (2009), and more recently summarized and further developed in Pham and Zelenyuk (2019). The method of Färe and Grosskopf (2003) is equivalent to that of Kuosmanen (2005) for CRS, yet they are generally different for VRS (see Pham and Zelenyuk, 2019 for recent theoretical results and related discussions).

In this study, we use the method in Kuosmanen (2005). We first partition Y_i as $Y_i = (G_i, B_i)$, where G_i and B_i are the desirable and undesirable (respectively) outputs in the sample. If $\mathcal{T}$ exhibits convexity and VRS, one could employ the VRS-DEA estimator given by

$$\widehat{\delta}_{\mathrm{VRS}}(x, g, b \mid d_x, d_g, d_b, \mathcal{X}_n) = \max_{\delta, s_1, \ldots, s_n, \alpha_1, \ldots, \alpha_n} \Big\{ \delta \geq 0 \mid (g + \delta d_g) \leqq \sum_{i=1}^{n} \alpha_i s_i G_i,$$
$$(b - \delta d_b) = \sum_{i=1}^{n} \alpha_i s_i B_i,\ (x - \delta d_x) \geqq \sum_{i=1}^{n} s_i X_i,\ \sum_{i=1}^{n} s_i = 1,\ \forall\, s_i \geq 0,\ \forall\, 0 \leq \alpha_i \leq 1 \Big\}, \tag{3.3}$$

which is a nonlinear programming (NLP) problem that can be reformulated as the following linear programming (LP) problem (Kuosmanen, 2005)[10]:

$$\widehat{\delta}_{\mathrm{VRS}}(x, g, b \mid d_x, d_g, d_b, \mathcal{X}_n) = \max_{\delta, \lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n} \Big\{ \delta \geq 0 \mid (g + \delta d_g) \leqq \sum_{i=1}^{n} \lambda_i G_i,$$
$$(b - \delta d_b) = \sum_{i=1}^{n} \lambda_i B_i,\ (x - \delta d_x) \geqq \sum_{i=1}^{n} (\lambda_i + \mu_i) X_i,\ \sum_{i=1}^{n} (\lambda_i + \mu_i) = 1,\ \forall\, \lambda_i,\ \mu_i \geq 0 \Big\}. \tag{3.4}$$

Alternatively, if $\mathcal{T}$ exhibits convexity and CRS, the CRS-DEA estimator could be used as follows,

$$\widehat{\delta}_{\mathrm{CRS}}(x, g, b \mid d_x, d_g, d_b, \mathcal{X}_n) = \max_{\delta, s_1, \ldots, s_n, \alpha_1, \ldots, \alpha_n} \Big\{ \delta \geq 0 \mid (g + \delta d_g) \leqq \sum_{i=1}^{n} \alpha_i s_i G_i,$$
$$(b - \delta d_b) = \sum_{i=1}^{n} \alpha_i s_i B_i,\ (x - \delta d_x) \geqq \sum_{i=1}^{n} s_i X_i,\ \forall\, s_i \geq 0,\ \forall\, 0 \leq \alpha_i \leq 1 \Big\}, \tag{3.5}$$

which can be reformulated as the following LP problem,

$$\widehat{\delta}_{\mathrm{CRS}}(x, g, b \mid d_x, d_g, d_b, \mathcal{X}_n) = \max_{\delta, \lambda_1, \ldots, \lambda_n} \Big\{ \delta \geq 0 \mid (g + \delta d_g) \leqq \sum_{i=1}^{n} \lambda_i G_i,$$
$$(b - \delta d_b) = \sum_{i=1}^{n} \lambda_i B_i,\ (x - \delta d_x) \geqq \sum_{i=1}^{n} \lambda_i X_i,\ \forall\, \lambda_i \geq 0 \Big\}, \tag{3.6}$$

where α_i disappeared because it can be shown its optimal value reached 1 for all i in instances of CRS, and so the programming problem becomes equivalent to the formulation in Färe and Grosskopf (2003).

To the best of our knowledge, the asymptotic properties for $\widehat{\delta}_{\mathrm{CRS}}(x, g, b \mid d_x, d_g, d_b, \mathcal{X}_n)$ and $\widehat{\delta}_{\mathrm{VRS}}(x, g, b \mid d_x, d_g, d_b, \mathcal{X}_n)$ are yet unknown in the literature. We fill the gap by deriving them formally in the following Section 4.

[10] To see this, let $\lambda_i = \alpha_i s_i$ and $\mu_i = s_i - \lambda_i = (1 - \alpha_i) s_i$. Using $0 \leq \alpha_i \leq 1$ and $0 \leq s_i \leq 1$, we can show that $0 \leq \lambda_i \leq 1$ and $0 \leq \mu_i \leq 1$. Substituting λ_i and μ_i into (3.3) yields (3.4).

4 Asymptotic Properties of the Estimators of Individual Efficiency: the Undesirable Outputs Case

4.1 Additional Assumptions

We concentrate on the presentation of the VRS case, because the CRS case follows analogous arguments. To get the result, we will need additional assumptions as in Kneip et al. (2008) and Simar et al. (2012), mainly:

Assumption 8. *Strict convexity of $\mathcal{T}$.*

Assumption 9. *Joint density for (X, G, B) on a compact support $\mathcal{D} \subseteq \mathcal{T}$.*

Assumption 10. *The observations in $\mathcal{X}_n$ are i.i.d on $\mathcal{T}$.*

Assumption 11. $\forall\ (x, g, b) \in \mathcal{T}$ *and* $\forall\ (d_x, d_g, d_b) \in \mathbb{R}_+^{p+q_g+q_b}$, $\delta(x, g, b|d_x, d_g, d_b, \mathcal{T})$ *is three times continuously differentiable.*

For the CRS case, Assumption 8 needs to be replaced by the following assumption.

Assumption 12. $\mathcal{T}$ *exhibits convexity and CRS.*

In what follows we will write $w = (x, g, b)$ as a generic point in $\mathcal{T}$ and we will use the notation $d = (-d_x, d_g, -d_b)$ where $d_x, d_g, d_b \geq 0$. Also we will use the simplified notation $\delta(w) = \delta(x, g, b|d_x, d_g, d_b, \mathcal{T})$ and $\widehat{\delta}_{\text{VRS}}(w) = \widehat{\delta}_{\text{VRS}}(x, g, b|d_x, d_g, d_b, \mathcal{X}_n)$ when there is no ambiguity.[11] Hence w has a size $r = p + q_g + q_b = p + q$.

4.2 Basic Arguments

We start as in Kneip et al. (2008) and Simar et al. (2012) and by looking to the directional distance of a particular production plan denoted by $w_0 = (x_0, g_0, b_0)$ and we analyse the DEA estimator of $\delta(w_0)$. This is given in equation (3.4) by replacing (x, g, b) by (x_0, g_0, b_0). In order to analyze the asymptotic properties of $\widehat{\delta}_{\text{VRS}}(w_0)$, we will alter the coordinate system to express both the frontier and its estimator in terms of scalar-valued functions. By doing this, we are able to rely on the basic results of Kneip et al. (2008) and Simar et al. (2012) to

[11] Note that $\delta(w)$ and $\widehat{\delta}_{\text{VRS}}(w)$ also depend on the direction, and are defined w.r.t. $\mathcal{T}$ and the sample $\mathcal{X}_n$, respectively, which we drop in the simplified notation to keep it simpler.

establish the statistical properties of the DEA estimators with the existence of undesirable outputs.

The transformation is in the spirit of the one suggested in Simar et al. (2012). Given a direction vector d, we can build $v_1, \ldots, v_{r-1}$, which is an orthonormal basis for d. We define the $r \times (r-1)$ matrix V with the jth column representing v_j. Clearly we have $V'd = \mathrm{O}_{r-1}$, a vector of zeros of dimension $r-1$. Simar et al. (2012) subsequently examine the linear transformation from $\mathbb{R}^r$ to $\mathbb{R}^r$ as follows:

$$h_{w_0} : w \to \xi = T(w - w_0), \tag{4.1}$$

where

$$T = \begin{pmatrix} V' \\ d'/||d|| \end{pmatrix}. \tag{4.2}$$

We see that T is a $r \times r$ orthogonal matrix, and Figure 1 in Simar et al. (2012) illustrates the transformation when $r = 2$. It is a translation $(w - w_0)$ to place the origin of the new axes at the point w_0 and the rotation T puts one coordinate, say u in the direction of d, and the remaining $r-1$ coordinates, say z, in a space orthogonal to d (and hence to u). If we define $\xi = (z'\ u)'$, we have

$$z = V'(w - w_0), \tag{4.3}$$

and

$$u = ||d||^{-1} d'(w - w_0). \tag{4.4}$$

Note that this is a one-to-one transformation, that can be easily inverted since $T'T = I_r$. We have

$$w = w_0 + T'\xi. \tag{4.5}$$

This transformation is specific to the chosen point w_0 and the chosen direction d, but can be used for all $w_0 \in \mathcal{T}$ and direction vector d. In this transformed coordinate system, we can depict the production set $\mathcal{T}$ as:

$$\Gamma_{w_0} = \{\xi \in \mathbb{R}^r \mid \xi = h_{w_0}(w), w \in \mathcal{T}\}. \tag{4.6}$$

For any point ξ (or w), u in (4.4) is a scalar measuring the gap between w and w_0 along the direction d (it is the projection of $w - w_0$ on d). So for this point ξ, the frontier will be reached by maximizing the value of u for this value of z, defining the scalar frontier function:

$$\phi_{w_0}(z) = \sup \{u \mid \xi = (z'\ u)' \in \Gamma_{w_0}\}. \tag{4.7}$$

For any arbitrary point $w \in \mathcal{T}$ and a given direction d, we also obtain

$$\delta(w) = ||d||^{-1}(\phi_{w_0}(z) - u), \tag{4.8}$$

where $\xi = (z' \ u)' = h_{w_0}(w)$. The latter can be seen by defining $w^\partial = w + \delta(w)d$, which is the frontier point for w. In the ξ units this point is given by $\xi^\partial = (z' \ \phi_{w_0}(z))'$. Simple calculation of $h_{w_0}(w^\partial)$ leads indeed to

$$\begin{aligned} \begin{pmatrix} z \\ \phi_{w_0}(z) \end{pmatrix} &= \begin{pmatrix} V'(w + \delta(w)d - w_0) \\ ||d||^{-1}d'(w + \delta(w)d - w_0) \end{pmatrix} \\ &= \begin{pmatrix} z \\ u + \delta(w)||d|| \end{pmatrix}, \end{aligned} \tag{4.9}$$

where for the last equality we used $V'd = \mathrm{O}_{r-1}$ and $d'd = ||d||^2$. Note that (4.8) provides what we wanted, i.e., a representation of the directional distance in the form of a scalar-valued function. Additionally, it should be noted that for the point of interest w_0, we have:

$$\delta(w_0) = ||d||^{-1}\phi_{w_0}(0), \tag{4.10}$$

because for w_0, both u and z are zero.

4.3 Asymptotic Properties

Now we will turn to the DEA estimator and see that even in the presence of undesirable outputs, which modify its definition, we keep a similar representation to the one derived for standard DEA estimators (as in Simar et al., 2012).

We first rewrite the DEA estimator (given by the linear program in equation (3.4)) as

$$\widehat{\delta}_{\mathrm{VRS}}(w) = \max_{\delta,\lambda,\mu,\beta} \{\delta \geq 0 \mid w + \delta d \in \widehat{\mathcal{T}}_{\mathrm{VRS}}\}, \tag{4.11}$$

where, due to the presence of undesirable outputs, $\widehat{\mathcal{T}}_{\mathrm{VRS}}$ is given by

$$\begin{aligned} \widehat{\mathcal{T}}_{\mathrm{VRS}} = \Bigg\{ (x, g, b) \mid x &= \sum_{i=1}^{n} (\lambda_i + \mu_i) X_i + \beta_x + \delta d_x, \\ g &= \sum_{i=1}^{n} \lambda_i G_i - \beta_g - \delta d_g, \\ b &= \sum_{i=1}^{n} \lambda_i B_i + \delta d_b, \\ i_n'(\lambda + \mu) &= 1, \lambda, \mu, \beta, \delta \geq 0 \Bigg\}, \end{aligned} \tag{4.12}$$

and where λ and μ are vectors of length n, $\beta = (\beta'_x \ \ \beta'_g)'$ has length $p + q_g$, and δ is scalar. The β's are there to ensure the strong disposability of the inputs and desirable outputs. In matrix notation, with $X = (X_1, \ldots, X_n) : p \times n$, $G = (G_1, \ldots, G_n) : q_g \times n$ and $B = (B_1, \ldots, B_n) : q_b \times n$ and $W = (X' \ G' \ B')' : r \times n$, we have

$$\widehat{\mathcal{T}}_{\text{VRS}} = \Bigg\{ w \ |w = W\lambda + \begin{pmatrix} X \\ \mathrm{O}_{q_g,n} \\ \mathrm{O}_{q_b,n} \end{pmatrix} \mu + \begin{pmatrix} \beta_x \\ -\beta_g \\ \mathrm{O}_{q_b,1} \end{pmatrix} - \delta d, \\ i'_n(\lambda + \mu) = 1, \lambda, \mu, \beta, \delta \geq 0 \Bigg\}. \tag{4.13}$$

In the new coordinates system defined above, this transforms into

$$\widehat{\Gamma}_{w_0,\text{VRS}} = \Bigg\{ \xi \ |\xi = TW\lambda + T \begin{pmatrix} X \\ \mathrm{O}_{q_g,n} \\ \mathrm{O}_{q_b,n} \end{pmatrix} \mu + T \begin{pmatrix} \beta_x \\ -\beta_g \\ \mathrm{O}_{q_b,1} \end{pmatrix} - T\delta d - Tw_0, \\ \lambda, \mu, \beta, \delta \geq 0, i'_n(\lambda + \mu) = 1 \Bigg\}, \tag{4.14}$$

which, by partitioning V as $V' = (V'_x \ V'_g \ V'_b)$ and remembering that $V'd = 0$ and $d'd = ||d||$, can be written element-wise as

$$\begin{aligned} &\widehat{\Gamma}_{w_0,\text{VRS}} = \Big\{ (z, u) \mid z = V'W\lambda + V'_x X\mu + V'_x \beta_x - V'_g \beta_g - V'w_0, \\ &u = ||d||^{-1} d'W\lambda - ||d||^{-1} d'_x X\mu - ||d||^{-1} d'_x \beta_x - ||d||^{-1} d'_g \beta_g - ||d||\delta - ||d||^{-1} d'w_0, \\ &i'_n(\lambda + \mu) = 1, \lambda, \mu, \beta, \delta \geq 0 \Big\}, \end{aligned} \tag{4.15}$$

Now, for a given z, the efficient boundary $\widehat{\Gamma}_{w_0,\text{VRS}}$ is given by the function

$$\widehat{\phi}_{w_0,\text{VRS}}(z) = \max_{u,\lambda,\mu,\beta,\delta} \{u \mid (z, u) \in \widehat{\Gamma}_{w_0,\text{VRS}}\}, \tag{4.16}$$

where we use the definition of $\widehat{\Gamma}_{w_0,\text{VRS}}$ described in (4.15). Since we maximize in u, and since $\delta \geq 0$ appears only in the definition of u with a negative sign, the solution of (4.16) in δ is $\delta = 0$. Thus, we arrive at the following definition of the function $\widehat{\phi}_{w_0,\text{VRS}}(z)$ for any z,

$$\begin{aligned} &\widehat{\phi}_{w_0,\text{VRS}}(z) = \max_{u,\lambda,\mu,\beta} \Big\{ u \mid z = V'W\lambda + V'_x X\mu + V'_x \beta_x - V'_g \beta_g - V'w_0, \\ &u = ||d||^{-1} d'W\lambda - ||d||^{-1} d'_x X\mu - ||d||^{-1} d'_x \beta_x - ||d||^{-1} d'_g \beta_g - ||d||^{-1} d'w_0, \\ &i'_n(\lambda + \mu) = 1, \lambda, \mu, \beta \geq 0 \Big\}, \end{aligned} \tag{4.17}$$

which is the analog to the equation (4.11) obtained in Simar et al. (2012), that had no undesirable outputs. The linear program is different but it has the same structure. Similarly for the true values of the DDF at any $w \in \mathcal{T}$, we have the VRS estimator

$$\widehat{\delta}_{\mathrm{VRS}}(w) = ||d||^{-1}(\widehat{\phi}_{w_0,\mathrm{VRS}}(z) - u), \tag{4.18}$$

where $\xi = (z'\ u)' = h_{w_0}(w)$. Moreover, at the specific point of interest w_0, we obtain

$$\widehat{\delta}_{\mathrm{VRS}}(w_0) = ||d||^{-1}\widehat{\phi}_{w_0,\mathrm{VRS}}(0). \tag{4.19}$$

As pointed out by Simar et al. (2012), the equations (4.10) and (4.19) indicate an equivalent formulation for defining the directional distances and their DEA estimator with the existence of undesirable outputs, similar to the approach used for the Farrell-type input efficiency in Kneip et al. (2008, 2015). Therefore, given the regularity conditions we obtain the desired results. First, for any point $w \in \mathcal{T}$ we have a nondegenerate limiting distribution with a rate of $n^{2/(p+q_g+q_b+1)} = n^{2/(p+q+1)}$ (as Theorem 1 in Kneip et al., 2008 and Theorem 4.1 in Simar et al., 2012). Secondly, we have the asymptotic behavior of the moments described for the Farrell-type input efficiency in Theorem 3.1 in Kneip et al. (2015). The following theorem provides the asymptotic moments for the VRS case.

Theorem 1. *Under Assumptions 1–4, 6–7, 8–11, as $n \to \infty$, a constant $0 < C_1 < \infty$ exists so that $\forall\ i \in \{1, 2, \ldots, n\}$,*

$$E\big(\widehat{\delta}_{VRS}(w_i) - \delta(w_i)\big) = C_1 n^{-\frac{2}{p+q_g+q_b+1}} + O\big(n^{-\frac{3}{p+q_g+q_b+1}}(\log n)^{\frac{p+q_g+q_b+4}{p+q_g+q_b+1}}\big), \tag{4.20}$$

and

$$VAR\big(\widehat{\delta}_{VRS}(w_i) - \delta(w_i)\big) = O\big(n^{-\frac{3}{p+q_g+q_b+1}}(\log n)^{\frac{3}{p+q_g+q_b+1}}\big), \tag{4.21}$$

and for $j \in \{1, 2, \ldots, i-1, i+1, \ldots, n\}$,

$$\Big|Cov\big(\widehat{\delta}_{VRS}(w_i) - \delta(w_i), \widehat{\delta}_{VRS}(w_j) - \delta(w_j)\big)\Big| = O\big(n^{-\frac{p+q_g+q_b+2}{p+q_g+q_b+1}}(\log n)^{\frac{p+q_g+q_b+2}{p+q_g+q_b+1}}\big) = o(n^{-1}). \tag{4.22}$$

For the CRS case, the asymptotic moments for any $w \in \mathcal{T}$ go along the same lines, as described in Park et al. (2010) and in Theorem 3.2 in Kneip et al. (2015). Under the CRS assumption, the rates are now given by $n^{2/(p+q_g+q_b)} = n^{2/(p+q)}$. The following theorem presents the asymptotic moments for the CRS case.

Theorem 2. *Under Assumptions 1–4, 6–7, 9–12, as $n \to \infty$, a constant $0 < C_2 < \infty$ exists so that $\forall\, i \in \{1, 2, \ldots, n\}$,*

$$E\big(\widehat{\delta}_{CRS}(w_i) - \delta(w_i)\big) = C_2 n^{-\frac{2}{p+q_g+q_b}} + O\big(n^{-\frac{3}{p+q_g+q_b}} (\log n)^{\frac{p+q_g+q_b+3}{p+q_g+q_b}}\big), \tag{4.23}$$

and

$$VAR\big(\widehat{\delta}_{CRS}(w_i) - \delta(w_i)\big) = O\big(n^{-\frac{3}{p+q_g+q_b}} (\log n)^{\frac{3}{p+q_g+q_b}}\big), \tag{4.24}$$

and for $j \in \{1, 2, \ldots, i-1, i+1, \ldots, n\}$,

$$\Big|Cov\big(\widehat{\delta}_{CRS}(w_i) - \delta(w_i), \widehat{\delta}_{CRS}(w_j) - \delta(w_j)\big)\Big| = O\big(n^{-\frac{p+q_g+q_b+1}{p+q_g+q_b}} (\log n)^{\frac{p+q_g+q_b+1}{p+q_g+q_b}}\big) = o(n^{-1}). \tag{4.25}$$

To summarize, the CRS and VRS estimators of $\delta(x, g, b \mid d_x, d_g, d_b, \mathcal{T})$ under weak disposability demonstrate consistency with a convergence rate of n^{κ}, where $\kappa = 2/(p + q_g + q_b) = 2/(p + q)$ if the CRS-DEA estimator is used, or $\kappa = 2/(p + q_g + q_b + 1) = 2/(p + q + 1)$ if the VRS-DEA estimator is applied.[12] It is clear that given the same dimensions, the estimators under weak disposability have the same convergence rate as the corresponding estimators under strong disposability. These results indicate that the aggregate DDFs under strong and weak disposability have the same theoretical results, which are discussed in the next Section 5.

5 CLTs and Inferences for Aggregate DDFs

Due to the results derived in the previous sections, the CLT results for the aggregate DDFs can be developed on the analogy with the aggregate Farrell-type output oriented measure of technical efficiency established by Simar and Zelenyuk (2018). We briefly summarize them here for future references and provide the guidance for practitioners who want to implement it by coding in a software or to modify it to certain needs. With a sample denoted as $\mathcal{X}_n = \{(X_i, Y_i)\}_{i=1}^n$, input price vector w_x, output price vector w_y, and the chosen set of directional vectors $\mathcal{D}_n = \{(d_x^i, d_y^i)\}_{i=1}^n$, we will focus on

$$\Delta_n = \sum_{i=1}^{n} \delta(X_i, Y_i \mid d_x^i, d_y^i, \mathcal{T}) S_d^i, \tag{5.1}$$

[12] The statistical results from Park et al. (2010) and Kneip et al. (2015) regarding the free disposal hull (FDH) estimator without undesirable outputs might be extended to the partial FDH estimator with undesirable outputs proposed by Podinovski and Kuosmanen (2011). However, the mathematical structure of the problem is completely different when using FDH approaches with undesirable outputs, necessitating further investigation. We defer this aspect to future research endeavors.

where S_d^i represents the weight of the ith DMU, given by

$$S_d^i = \frac{w_y d_y^i + w_x d_x^i}{\sum_{i=1}^n (w_y d_y^i + w_x d_x^i)}. \tag{5.2}$$

To simplify the notation, denote $H_{1i} = \delta(X_i, Y_i \mid d_x^i, d_y^i, \mathcal{T}) H_{2i}$, where $H_{2i} = w_y d_y^i + w_x d_x^i$. Then equation (5.1) becomes

$$\Delta_n = \frac{n^{-1} \sum_{i=1}^n H_{1i}}{n^{-1} \sum_{i=1}^n H_{2i}}, \tag{5.3}$$

which is an estimate of the ratio of the true means

$$\Delta = \frac{E(H_{1i})}{E(H_{2i})} = \frac{\mu_1}{\mu_2}. \tag{5.4}$$

If we observe the true DDF, $\delta(X_i, Y_i \mid d_x^i, d_y^i, \mathcal{T})$, according to the typical CLT for the ratio of means (Kendall and Stuart, 1968), the following holds:

$$\sqrt{n}\,(\Delta_n - \Delta) \xrightarrow{\mathcal{L}} N(0, \sigma_\Delta^2), \tag{5.5}$$

where

$$\sigma_\Delta^2 = \Delta^2 \left[\frac{\sigma_1^2}{\mu_1^2} + \frac{\sigma_2^2}{\mu_2^2} - 2\frac{\sigma_{12}}{\mu_1 \mu_2} \right]. \tag{5.6}$$

In equation (5.6), $\sigma_1^2 = \mathrm{Var}(H_{1i})$, $\sigma_2^2 = \mathrm{Var}(H_{2i})$, and $\sigma_{12} = \mathrm{Cov}(H_{1i}, H_{2i})$.[13] However, as $\delta(X_i, Y_i \mid d_x^i, d_y^i, \mathcal{T})$ is not directly observed, the corresponding Δ_n remains unobserved. Our best recourse is to derive an estimate of $\delta(X_i, Y_i \mid d_x^i, d_y^i, \mathcal{T})$ from $\mathcal{X}_n$, for instance, through methods like CRS-DEA, or VRS-DEA as elaborated in Subsection 3.1 for the standard case or Subsection 3.2 for undesirable outputs case, to obtain $\widehat{\delta}(X_i, Y_i \mid d_x^i, d_y^i, \mathcal{X}_n)$.

Substitute $\widehat{\delta}(X_i, Y_i \mid d_x^i, d_y^i, \mathcal{X}_n)$ into (5.3) to obtain an estimate of Δ_n, given by:

$$\widehat{\Delta}_n = \frac{\widehat{\mu}_{1,n}}{\widehat{\mu}_{2,n}} = \frac{n^{-1} \sum_{i=1}^n \widehat{H}_{1i}}{n^{-1} \sum_{i=1}^n H_{2i}}, \tag{5.7}$$

where $\widehat{H}_{1i} = \widehat{\delta}(X_i, Y_i \mid d_x^i, d_y^i, \mathcal{X}_n) H_{2i}$, is the estimate of H_{1i}. The following theorem formulates the basic properties for $\widehat{\Delta}_n$.

[13] As discussed in Section 2, when we have the fixed direction, i.e., the directional vector $(d_x^i, d_y^i) = (d_x, d_y)$ for all i, the aggregate DDF Δ_n reduces to the simple mean DDF. However, the CLT for Δ_n in (5.5) is still valid for the simple mean DDF. This is left as an exercise for the reader.

Theorem 3. *Subject to the relevant set of assumptions detailed in the preceding sections for the case of CRS-DEA or VRS-DEA, we have*

$$E(\widehat{\Delta}_n) = \Delta + Cn^{-\kappa} + o(n^{-\kappa}), \tag{5.8}$$

$$\widehat{\Delta}_n - E(\widehat{\Delta}_n) = \Delta_n - \Delta + o_p(n^{-1/2}), \tag{5.9}$$

$$\sqrt{n}\left(\widehat{\Delta}_n - E(\widehat{\Delta}_n)\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2_\Delta), \tag{5.10}$$

where C is an unknown constant term.

From (5.8), we see that the estimator $\widehat{\Delta}_n$ consistently estimates Δ, albeit with a bias term on the order of $n^{-\kappa}$. The above theorem also shows why the standard CLT cannot work. To see this, substituting (5.8) into (5.10) yields

$$\sqrt{n}\left(\widehat{\Delta}_n - \Delta - Cn^{-\kappa} - o(n^{-\kappa})\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2_\Delta). \tag{5.11}$$

The bias term $Cn^{-\kappa}$ multiplied by $\sqrt{n}$ is constant if $\kappa = 1/2$, or explodes if $\kappa < 1/2$; in these two cases, the standard CLT which ignores the bias term $Cn^{-\kappa}$ cannot be used. Otherwise (i.e., if $\kappa > 1/2$), it can be ignored and the standard CLT can be applied. This result indicates that in the presence of undesirable outputs, the standard CLT cannot work in general except in one CRS case (i.e., when $p = 1$, $q_g = 1$, and $q_b = 1$ so that $\kappa = 2/3$). Now we can formulate the new CLTs for $\widehat{\Delta}_n$. The essence of such developments is presented in the following theorem.

Theorem 4. *Subject to the relevant set of assumptions detailed in the preceding sections for the case of CRS-DEA or VRS-DEA, if $\kappa \geq 1/2$, we obtain*

$$\sqrt{n}\left(\widehat{\Delta}_n - \widehat{\mathcal{B}}_{\Delta,n} - \Delta - o(n^{-\kappa})\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2_\Delta); \tag{5.12}$$

and if $\kappa < 1/2$, we have

$$\sqrt{n_\kappa}\left(\widehat{\Delta}_{n_\kappa} - \widehat{\mathcal{B}}_{\Delta,n} - \Delta - o(n^{-\kappa})\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2_\Delta), \tag{5.13}$$

where

$$\kappa = \begin{cases} 2/(p+q), & \text{if the CRS-DEA estimator is used,} \\ 2/(p+q+1), & \text{if the VRS-DEA estimator is used.} \end{cases} \tag{5.14}$$

In Theorem 4, $\widehat{\mathcal{B}}_{\Delta,n}$ is the bias estimate for $\widehat{\Delta}_n$, akin to the approach taken by Simar and Zelenyuk (2018) for the aggregate Farrell-type output oriented measure of technical efficiency.[14] Moreover, $\widehat{\Delta}_{n_\kappa}$ is a subsample variant of $\widehat{\Delta}_n$, with a size of $n_\kappa = \lfloor n^{2\kappa} \rfloor < n$ rather than n.[15] More specifically,

$$\widehat{\Delta}_{n_\kappa} = \frac{n_\kappa^{-1} \sum_{\{i|(X_i,Y_i)\in\mathcal{S}_{n_\kappa}\}} \widehat{\delta}(X_i, Y_i \mid d_x^i, d_y^i, \mathcal{X}_n) H_{2i}}{n_\kappa^{-1} \sum_{\{i|(X_i,Y_i)\in\mathcal{S}_{n_\kappa}\}} H_{2i}}, \tag{5.15}$$

where $\mathcal{S}_{n_\kappa}$ represents a randomized subset with a sample size of n_κ, derived from $\mathcal{S}_n$.[16]

Theorem 4 could enable researchers to conduct the statistical inference. Specifically, the asymptotic confidence intervals with a level of $100(1-\alpha)\%$ for the true aggregate Δ derived from Theorem 4 are given by

$$\left[\widehat{\Delta}_n - \widehat{\mathcal{B}}_{\Delta,n} \ \pm \ \Phi^{-1}_{1-\alpha/2}\, \widehat{\sigma}_\Delta / \sqrt{n}\right], \text{ if } \kappa \geq 1/2, \tag{5.16}$$

and

$$\left[\widehat{\Delta}_{n_\kappa} - \widehat{\mathcal{B}}_{\Delta,n} \ \pm \ \Phi^{-1}_{1-\alpha/2}\, \widehat{\sigma}_\Delta / \sqrt{n_\kappa}\right], \text{ if } \kappa < 1/2, \tag{5.17}$$

where $\widehat{\sigma}_\Delta$ is an estimate of σ_Δ, given by

$$\widehat{\sigma}^2_{\Delta,n} = \widehat{\Delta}^2_n \left[\frac{\widehat{\sigma}^2_{1,n}}{\widehat{\mu}^2_{1,n}} + \frac{\widehat{\sigma}^2_{2,n}}{\widehat{\mu}^2_{2,n}} - 2\frac{\widehat{\sigma}_{12,n}}{\widehat{\mu}_{1,n}\widehat{\mu}_{2,n}}\right], \tag{5.18}$$

and where $\widehat{\sigma}^2_{1,n} = \widehat{\mathrm{Var}}(\widehat{H}_{1i})$, $\widehat{\sigma}^2_{2,n} = \widehat{\mathrm{Var}}(H_{2i})$, and $\widehat{\sigma}_{12,n} = \widehat{\mathrm{Cov}}(\widehat{H}_{1i}, H_{2i})$.[17]

To conclude this section, it is worth noting that while the theory presented so far and the simulations that follow are done specifically for the case when CRS-DEA and VRS-DEA estimators are used, they should also be valuable as a guide for adaptations to the contexts when other methods are used for estimating the directional distance functions and their aggregates.

[14] The procedures are described more precisely in the separate Appendix A of the online Supplement.

[15] $\lfloor n^{2\kappa} \rfloor$ indicates the largest integer that is no greater than $n^{2\kappa}$.

[16] As pointed out in the Remark 1 of Simar and Zelenyuk (2018), when $\kappa = 2/5$ in the case of CRS-DEA or VRS-DEA, both (5.12) and (5.13) can be used. However, it is advisable to use (5.13) due to a smaller remainder term, which is also evident in the simulations presented in the next section.

[17] There are several other options in the literature to estimate σ_Δ. In particular, one can consider to adapt the alternative estimators for the variance of the simple mean and aggregate technical efficiency proposed by Simar and Zelenyuk (2020) and Simar et al. (2023a, 2024) and the data sharpening method recommended by Nguyen et al. (2022) to the case of the aggregate DDFs. These adaptions to the DDF case are not obvious and may need further examinations, which we leave for further research.

6 Monte-Carlo Evidence

To our knowledge, the existing literature lacks clarity on the best ways to *parametrically* characterize technologies with undesirable outputs that are cohesive with the existing DEA estimators, either in theory or in simulations and so in this section we will attempt to explore and shed light on this crucial aspect of modeling. There are different ways to think of how one should view the true (or hypothetical) technology that involves undesirable outputs and, perhaps, there is no one specific way that fits all (or most) realistic scenarios, as it might depend on the context.

For example, one could think of the undesirable output as being a function (e.g., monotonically increasing) or more generally as a correspondence of the desirable output given inputs. In other words, a firm might be able to use some inputs to produce a bundled output composed of both the desirable and undesirable outputs, where reducing the undesirable output may require reducing the desirable output. An example of such relationships could be seen in the banking industry: the undesirable output (e.g., the non-performing loans) is often viewed/expected as a function (related to probability of success/failure) of the desirable output (e.g., good loans).

Alternatively, one could think of an undesirable output as being a function (or a correspondence) of some inputs, which are used in a production of (and hence enters into a function or a correspondence with) the desirable output, along with other inputs. An example of such relationships could be seen in the energy generation business: the undesirable output (pollution) is often viewed as being a function of a particular input (e.g., coal), defined by some natural laws, which may be substituted by another input (e.g., gas) with a different functional relationship with the same pollution or yet another input (solar) with no (known) relationship with the pollution.

We will explore both of these scenarios here. Hereafter, we will provide the MC results for the first scenario while leaving the results for the second scenario to the separate Appendix B of the online Supplement.

6.1 The Data Generation Process

For the simulations in the main text, we consider the following technology

$$\mathcal{Y}_i^\partial = c\prod_{j=1}^{p}(x_{ij}^\partial)^{\beta_j}, \tag{6.1}$$

where c is a constant term, x_{ij}^∂ is the jth efficient input, and $\mathcal{Y}_i^\partial$ is the efficient bundled output of desirable and undesirable outputs, specified as,

$$\mathcal{Y}_i^\partial = g_i^\partial/(b_i^\partial)^{\beta_b}, \tag{6.2}$$

and where g_i^∂ and b_i^∂ are the desirable and undesirable (respectively) outputs on the frontier. The technologies specified in (6.1) and (6.2) indicate that given fixed inputs, to reduce the undesirable outputs, it is necessary to decrease the desirable outputs, i.e., Assumption 7 is satisfied.

In the MC experiments, we let the specified directional vector be the observed inputs and outputs, i.e., $(d_x^i, d_g^i, d_b^i) = (x_i, g_i, b_i)$, which means the directional distance $\delta_i = \delta(x_i, g_i, b_i \mid x_i, g_i, b_i, \mathcal{T}) \in [0,1]$ measures the maximal percentage of feasible increase of the desirable outputs and simultaneous decrease (with the same percentage δ_i) of the undesirable outputs and inputs. The ith true directional distance score is generated from $\delta_i \overset{\text{iid}}{\sim} \text{Beta}(1,2)$, i.e., we generate δ_i from the Beta distribution to ensure $\delta_i \in [0,1]$ for all $i = 1,\ldots,n$. In addition, we generate the ith observed input as $x_i \overset{\text{iid}}{\sim} \text{Unif}(0,1)$ and observed undesirable output as $b_i \overset{\text{iid}}{\sim} \text{Unif}(0,1)$, then calculate the points on the frontier as $(x_i^\partial, g_i^\partial, b_i^\partial)$, where $x_i^\partial = x_i(1-\delta)$, $b_i^\partial = b_i(1-\delta)$, and $g_i^\partial = c(b_i^\partial)^{\beta_b}\prod_{j=1}^{p}(x_{ij}^\partial)^{\beta_j}$. Finally, we set the ith observed output as $g_i = g_i^\partial/(1+\delta_i)$. Thus, we obtain a simulated sample $\mathcal{X}_n = \{(x_i, g_i, b_i)\}_{i=1}^n$. As the technology exhibits VRS, we employ the VRS-DEA estimator in the simulations.

In each experiment, 1,000 MC trials are performed. During each trial, we assess the coverage, i.e., whether the true value of the aggregate DDFs is covered by the estimated confidence intervals. Subsequently, the average coverage is reported across these 1,000 MC trials. The true values for the aggregate DDF are calculated through using the randomly simulated 10,000,000 realizations of δ_i, x_i, g_i, b_i and the specified prices (w_x, w_g, w_b). Table 1 presents the parameter values used in the MC experiments.

Before showcasing our Monte Carlo results, for ease of notation, we designate as follows: (i) the standard CLT (i.e., using the CLT in (5.5) by replacing Δ_n and σ_Δ^2 by $\widehat{\Delta}_n$ in (5.7)

Table 1: The Values of Parameters in the MC

p	1	2	3	4
β_1	0.4	0.4	0.4	0.4
β_2		0.2	0.2	0.2
β_3			0.1	0.1
β_4				0.15
β_b	0.2	0.1	0.15	0.05
c	1	1	1	1
w_1	1	1	1	1
w_2		2	2	2
w_3			0.5	0.5
w_4				1
w_g	1	1	1	1
w_b	1	1	1	1

and $\widehat{\sigma}_{\Delta}^2$ in (5.18), respectively), which ignores the bias correction and uses the wrong rate of convergence.[18]

(ii) the developed CLT (with bias correction) for the aggregate DDFs with n observations (i.e., using (5.16));

(iii) the developed CLT (with bias correction) for the aggregate DDFs with n_κ observations (i.e., using (5.17)).

6.2 Main Results

The simulation results are reported in Table 2. Panel A presents the results with low dimensions where $p + q_g + q_b < 4$ while Panel B presents those for high dimensions where $p + q_g + q_b \geq 4$. From Panel A, we observe that with the sample size increasing, the coverages using approach (ii) for the aggregate DDFs with the existence of undesirable outputs converge towards the corresponding nominal coverages. In contrast, those using approach (i) (i.e., the standard CLT) fail to provide any sensible results, as expected by the theory: the bias will explode and the rate is incorrect. This is also true when we examine the re-

[18] Further insights into why the standard CLT cannot work in this case are in the discussions following Theorem 3.

Table 2: Coverages of Estimated Confidence Intervals for Aggregate Directional Distance Functions using VRS-DEA Estimator

Panel A: The Cases with $p + q_g + q_b < 4$

				0.90		0.95		0.99	
p	q_g	q_b	n	(i)	(ii)	(i)	(ii)	(i)	(ii)
1	1	1	10	0.144	0.501	0.216	0.580	0.360	0.690
1	1	1	20	0.124	0.647	0.198	0.720	0.352	0.842
1	1	1	50	0.105	0.747	0.170	0.820	0.345	0.934
1	1	1	100	0.111	0.817	0.182	0.888	0.397	0.958
1	1	1	200	0.123	0.839	0.182	0.911	0.388	0.968
1	1	1	500	0.114	0.856	0.182	0.918	0.382	0.972
1	1	1	1000	0.119	0.874	0.187	0.934	0.388	0.983

Panel B: The Cases with $p + q_g + q_b \geq 4$

				0.90			0.95			0.99		
p	q_g	q_b	n	(i)	(ii)	(iii)	(i)	(ii)	(iii)	(i)	(ii)	(iii)
2	1	1	10	0.032	0.324	0.386	0.065	0.376	0.445	0.146	0.451	0.527
2	1	1	20	0.020	0.491	0.592	0.039	0.549	0.670	0.104	0.669	0.776
2	1	1	50	0.010	0.631	0.716	0.015	0.712	0.796	0.042	0.825	0.908
2	1	1	100	0.000	0.726	0.827	0.002	0.795	0.895	0.022	0.911	0.967
2	1	1	200	0.000	0.746	0.837	0.001	0.823	0.903	0.015	0.927	0.966
2	1	1	500	0.000	0.814	0.865	0.000	0.891	0.926	0.002	0.963	0.982
2	1	1	1000	0.000	0.841	0.895	0.000	0.903	0.947	0.002	0.965	0.985
3	1	1	10	0.005	0.176	0.253	0.012	0.205	0.309	0.029	0.274	0.381
3	1	1	20	0.001	0.272	0.414	0.003	0.319	0.483	0.012	0.404	0.602
3	1	1	50	0.000	0.460	0.680	0.000	0.533	0.778	0.002	0.665	0.891
3	1	1	100	0.000	0.566	0.775	0.000	0.642	0.859	0.000	0.770	0.943
3	1	1	200	0.000	0.654	0.856	0.000	0.742	0.916	0.000	0.869	0.973
3	1	1	500	0.000	0.713	0.874	0.000	0.793	0.920	0.000	0.921	0.980
3	1	1	1000	0.000	0.726	0.884	0.000	0.810	0.937	0.000	0.915	0.985
4	1	1	10	0.002	0.081	0.151	0.003	0.102	0.182	0.006	0.141	0.233
4	1	1	20	0.000	0.127	0.291	0.000	0.158	0.353	0.000	0.227	0.448
4	1	1	50	0.000	0.255	0.523	0.000	0.308	0.605	0.000	0.404	0.753
4	1	1	100	0.000	0.392	0.730	0.000	0.456	0.811	0.000	0.556	0.911
4	1	1	200	0.000	0.528	0.822	0.000	0.600	0.909	0.000	0.722	0.972
4	1	1	500	0.000	0.639	0.871	0.000	0.715	0.928	0.000	0.818	0.983
4	1	1	1000	0.000	0.598	0.875	0.000	0.692	0.937	0.000	0.817	0.983

sults in Panel B for high dimensions. The coverages using approach (iii) increase toward the nominal coverages while those using approach (i) go to zero when the sample size increases. These results validate our developed theoretical results for the aggregate DDFs in Section 5. Moreover, comparing coverages using approaches (ii) and (iii) in Panel B, we see that the performance of approach (ii) is not so good (far away from the nominal coverage) when $p = 3$ or 4. When $p = 2$, as n increases, the coverages using approach (ii) also increase toward to the nominal coverage, but are much smaller than that using approach (iii), which verifies the previous arguments that if $\kappa = 2/5$ for VRS-DEA case, both (5.12) and (5.13) can be used. However (5.13) is recommended.

Moreover, it should be noted that when the sample size is not large (e.g., ≤ 500), the estimated confidence intervals based on (ii) or (iii) undercover the true values. For example, when the sample size is 100 and $p = 2$, the coverage using (iii) is only 0.827 when the nominal coverage is 0.900. Further, this undercovering phenomenon is more severe when the dimensions $(p + q_g + q_b)$ increase. This undercovering phenomenon in large dimensions and finite sample sizes is well known for the various nonparametric efficiency estimators, e.g., see also the related discussions in Simar and Zelenyuk (2020), Nguyen et al. (2022), and Simar et al. (2023a, 2024).

7 An Empirical Illustration

To illustrate our developed theoretical results for the aggregate DDFs, we use the data set for the large US commercial banks obtained from the Federal Reserve Bank of Chicago.[19] This data set was used by Malikov et al. (2016) and more recently by Tsionas et al. (2023) in the SFA context. Following Malikov et al. (2016) and Wheelock and Wilson (2018), five desirable outputs, five inputs and one undesirable output (the non-performing loans, b_i) are specified for the banks' production function. Theorem 4 shows that given the sample size, the estimation errors for the non-parametric efficiency estimators increase with the dimensions. To enhance the accuracy of the estimation, we employ the dimension reduction method introduced by Daraio and Simar (2007) to consolidate five desirable outputs into one

[19] The code for this illustration can be found on GitHub at the following link: `https://github.com/srzhao89/szz-ddf`.

Table 3: VRS-DEA Estimates of Aggregate DDFs, Standard Deviations, and the 95% Confidence Intervals for the Large US Banks

					95% CI			
Year	$\widehat{\Delta}_n$	$\widehat{B}_{\Delta,n}$	$\widehat{\Delta}_n - \widehat{B}_{\Delta,n}$	$\widehat{\sigma}_\Delta$	(i)		(ii)	
2001	0.0706	-0.0506	0.1213	0.2552	0.0404	0.1009	0.0911	0.1515
2002	0.0574	-0.0381	0.0954	0.2307	0.0302	0.0845	0.0683	0.1226
2003	0.0670	-0.0434	0.1105	0.2595	0.0364	0.0977	0.0798	0.1411
2004	0.0558	-0.0308	0.0865	0.2346	0.0281	0.0834	0.0589	0.1142
2005	0.0556	-0.0350	0.0907	0.2356	0.0271	0.0841	0.0622	0.1191
2006	0.0534	-0.0388	0.0922	0.2217	0.0256	0.0811	0.0644	0.1199
2007	0.0494	-0.0390	0.0884	0.2098	0.0224	0.0764	0.0614	0.1154
2008	0.0336	-0.0236	0.0572	0.1530	0.0119	0.0553	0.0355	0.0789
2009	0.0292	-0.0166	0.0458	0.1305	0.0103	0.0480	0.0269	0.0646
2010	0.0426	-0.0321	0.0746	0.1998	0.0133	0.0718	0.0454	0.1039

NOTE: (i) means using the standard CLT; (ii) means relying on our developed CLT, i.e., using (5.16).

aggregated output g_i^*, and five inputs to one aggregated input x_i^*.[20] Through this process, we obtain the aggregated input-output data, $\mathcal{X}_n = \{(x_i^*, g_i^*, b_i)\}_{i=1}^n$, where $p = 1$, $q_g = 1$ and $q_b = 1$.[21]

In addition, for the illustration purpose, we set the price of the non-performing loans as -1. Note that the input prices are available while the output prices are not provided by Malikov et al. (2016). Hence, we choose $(d_{x^*}^i, d_{g^*}^i, d_b^i) = (x_i^*, 0, b_i)$ when estimating the aggregate DDFs, which depends on prices and specified directions as expressed in (2.7). By doing so, the individual weight S_d^i in (2.8) does not depend on the desirable outputs or their prices. In this situation the estimate of the aggregate DDFs under weak disposibility measures the aggregate (or the weighted-average) of the individual maximal percentages of feasible decrease of the aggregated input and the non-performing loans while maintaining the aggregated desirable output and technology unchanged.

Table 3 summarizes the estimation outcomes obtained through VRS-DEA estimators

[20] Wilson (2018) has shown in various MC experiments that the gain of accuracy in statistical inference can be important by using this dimension reduction method.

[21] For more details about the specifications of inputs and outputs and also the dimension reduction method, see the separate Appendix C of the online Supplement.

when the non-performing loans are considered as undesirable outputs and modeled under weak disposibility, utilizing the estimator outlined in (3.4). It is worth noting that modeling the undesirable outputs in the framework of weak disposability is more consistent with economic theory and also reflects the production for the firms in practice.

First, it is beneficial to reiterate the intuitive meaning of efficiency based on the DDF. For example, the estimate of the true DDF for the bank with the ID equal to 89 is 0.0784 in 2001, indicating that this bank has the potential to simultaneously decrease the aggregated input and the undesirable output by 7.84 percent while keeping the aggregated desirable output and technology constant, if it reaches the estimated best-practice frontier, in the chosen direction of measurement (i.e., $(-x_i^*, 0, -b_i)$ here, where x_i^* and b_i are the values observed for this bank).[22] We then aggregate by employing a specific formula that computes the weighted average of these DDFs across all observations, where the weights reflect an economic significance corresponding to the measurement orientation, according to (2.19) and (2.20). By doing so, we obtain the estimate of the aggregate DDFs before the bias correction. Then, we implement the bias correction procedure outlined in the separate Appendix A of the online Supplement, estimate the standard errors using (5.18), and obtain the confidence intervals using (5.16).

Second, we see that the differences among the estimates of the aggregate DDFs before and after the bias correction are very large for most years. For example, the estimate of the aggregate DDFs after the bias correction in 2001 (0.1213) is more than 1.7 times larger than that before the bias correction (0.0706), thus illustrating the importance in the empirical research to correct the bias for the estimate of the aggregate DDFs.

Third, the 95% CIs constructed using approach (ii) (i.e., the developed CLT with $\kappa \geq 1/2$ as $p + q_g + q_b = 3$) exhibit significant differences from that using approach (i) (i.e., the standard CLT) for each year. As indicated by our MC results in Section 6, the standard CLT demonstrates poor performance for the aggregate DDFs, implying that the statistical inference relying on the standard CLT in this context might lead to misleading results. For example, approach (ii) indicates that the true value of the aggregate DDFs for the large US banks in 2001 was larger than 0.0706 (i.e., the estimate of the aggregate DDFs before the

[22] For the sake of conserving the space, the individual estimates are not provided in the table but are available upon request.

bias correction) as the lower bound of the 95% CIs constructed using approach (ii) was larger than 0.0706. On the other hand, approach (i) cannot reject the null hypothesis that the true value is 0.0706 as the 95% CIs constructed using approach (i) contain 0.0706 in 2001.

Finally, the large US banks became more efficient (relative to their annual best-practice frontiers) over time as a decreasing trend is observed for the bias-corrected estimates of the aggregate DDFs. For example, Panel A shows that the estimates of the aggregate DDFs for the large US banks have gradually (with some fluctuations) decreased from 0.1213 in 2001 to 0.0746 in 2010. Moreover, the large US banks were more efficient especially after 2007, indicating that although the global financial crisis has greatly disrupted the banking sector, it might have also brought a benefit by enforcing the large US banks to improve their performance to get through these tough times.

8 Conclusions

This article contributes significantly to the existing literature in two key aspects. On one hand, we established the central limit theorems for aggregate directional distance functions, which embed those for the simple mean and aggregate Farrell-type efficiency as special cases. Consequently, our paper supplements to the recent developments of the statistical results for the various nonparametric efficiency estimators, such as the Farrell-type technical efficiency (Kneip et al., 2015; Simar and Zelenyuk, 2018), Malmquist productivity indices (Pham et al., 2024; Kneip et al., 2021), Hicks–Moorsteen productivity indices (Simar et al., 2023b), etc.

On the other hand and most importantly, we develop the central limit theorems for aggregate directional distance functions under the weak disposability of undesirable and desirable outputs and also conduct Monte-Carlo experiments to assess the efficacy of the developed central limit theorems. The simulations affirm that employing the developed central limit theorems for statistical inference yields significantly improved performance in comparison to the standard central limit theorems. As far as we are aware, the statistical results of the various non-parametric efficiency methods under weak disposability have not been investigated, either theoretically or through simulations, and so we endeavor to make a breakthrough on this important aspect of modeling. We also use the data from the large US banks to illustrate the developed theoretical results for the aggregate directional distance functions in the presence of undesirable outputs. The results imply that utilizing the traditional CLTs for statistical inference in this context might lead to potentially misleading outcomes.

Future research could try to adapt the recently introduced improving methods from Simar and Zelenyuk (2020), Nguyen et al. (2022), Zelenyuk and Zhao (2023), and Simar et al. (2023a, 2024) to further increase the performance of the statistical results for the aggregate DDFs with general disposability. Another fruitful avenue for further research could be to derive new CLTs for the aggregate productivity indexes based on the DDFs (for technologies with undesirable outputs and accounting for economic weights), and their decompositions. Such work can build on the earlier works paved by Jeon and Sickles (2004) and more recent CLTs established by Simar and Wilson (2019), Kneip et al. (2021) and Pham et al. (2024).

Declarations

Acknowledgement

Valentin Zelenyuk acknowledges the support from the Australian Research Council (FT170100401) and The University of Queensland. Shirong Zhao acknowledges the Liaoning Provincial Department of Education Research Fund (LJKMZ20221602). Feedback from Zhichao Wang and Evelyn Smart is appreciated. We also thank Paul Wilson as the programming codes used in this paper involve some earlier codes learned from Paul Wilson. These individuals and organizations are not responsible for the views expressed here. These authors contributed equally: Léopold Simar, Valentin Zelenyuk, Shirong Zhao.

Conflict of interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and publication of this article.

References

Aigner, D., C. A. K. Lovell, and P. Schmidt (1977), Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics* 6, 21–37.

Banker, R. D., A. Charnes, and W. W. Cooper (1984), Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science* 30, 1078–1092.

Bogetoft, P. (1994), Incentive efficient production frontiers: An agency perspective on DEA, *Management Science* 40, 959–968.

Centorrino, S. and C. F. Parmeter (2024), Nonparametric estimation of stochastic frontier models with weak separability, *Journal of Econometrics* 238, 105641.

Chambers, R. G., Y. Chung, and R. Färe (1996), Benefit and distance functions, *Journal of Economic Theory* 70, 407–419.

— (1998), Profit, directional distance functions, and nerlovian efficiency, *Journal of Optimization Theory and Applications* 98, 351–364.

Charnes, A., W. W. Cooper, and E. Rhodes (1978), Measuring the efficiency of decision making units, *European Journal of Operational Research* 2, 429–444.

Daraio, C. and L. Simar (2007), *Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications*, New York: Springer Science & Business Media.

Deprins, D., L. Simar, and H. Tulkens (1984), Measuring labor inefficiency in post offices, in M. M. P. Pestieau and H. Tulkens, eds., *The Performance of Public Enterprises: Concepts and Measurements*, Amsterdam: North-Holland, pp. 243–267.

Färe, R., S. Grosskopf, and C. A. K. Lovell (1994), *Production Frontiers*, Cambridge: Cambridge University Press.

Färe, R., S. Grosskopf, C. A. K. Lovell, and C. Pasurka (1989), Multilateral productivity comparisons when some outputs are undesirable: A nonparametric approach, *Review of Economics and Statistics* 71, 90–98.

Färe, R., S. Grosskopf, C. K. Lovell, and S. Yaisawarng (1993), Derivation of shadow prices for undesirable outputs: A distance function approach, *Review of Economics and Statistics* 75, 374–380.

Färe, R., S. Grosskopf, D.-W. Noh, and W. Weber (2005), Characteristics of a polluting technology: Theory and practice, *Journal of Econometrics* 126, 469–492.

Färe, R., S. Grosskopf, and W. L. Weber (2006), Shadow prices and pollution costs in US agriculture, *Ecological Economics* 56, 89–103.

Färe, R. and D. Primont (1995), *Multi-output Production and Duality: Theory and Applications*, Dordrecht: Kluwer Academic Publishers.

Farrell, M. J. (1957), The measurement of productive efficiency, *Journal of the Royal Statistical Society A* 120, 253–281.

Färe, R. and S. Grosskopf (2003), Nonparametric productivity analysis with undesirable outputs: Comment, *American Journal of Agricultural Economics* 85, 1070–1074.

— (2009), A comment on weak disposability in nonparametric production analysis, *American Journal of Agricultural Economics* 91, 535–538.

Färe, R. and V. Zelenyuk (2003), On aggregate Farrell efficiencies, *European Journal of Operational Research* 146, 615–620.

Greene, W. (2005), Reconsidering heterogeneity in panel data estimators of the stochastic frontier model, *Journal of Econometrics* 126, 269–303.

Jeon, B. M. and R. C. Sickles (2004), The role of environmental factors in growth accounting, *Journal of Applied Econometrics* 19, 567–591.

Kendall, M. and A. Stuart (1968), *The Advanced Theory of Statistics*, Charles Griffin & Company, London, UK.

Kneip, A., L. Simar, and P. W. Wilson (2008), Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models, *Econometric Theory* 24, 1663–1697.

— (2015), When bias kills the variance: Central limit theorems for DEA and FDH efficiency scores, *Econometric Theory* 31, 394–422.

— (2021), Inference in dynamic, nonparametric models of production: Central limit theorems for malmquist indices, *Econometric Theory* 37, 537–572.

Korostelev, A., L. Simar, and A. B. Tsybakov (1995), Efficient estimation of monotone boundaries, *The Annals of Statistics* 23, 476–489.

Kuosmanen, T. (2005), Weak disposability in nonparametric production analysis with undesirable outputs, *American Journal of Agricultural Economics* 87, 1077–1082.

Kuosmanen, T. and V. Podinovski (2009), Weak disposability in nonparametric production analysis: Reply to Färe and Grosskopf, *American Journal of Agricultural Economics* 91, 539–545.

Malikov, E., S. C. Kumbhakar, and M. G. Tsionas (2016), A cost system approach to the stochastic directional technology distance function with undesirable outputs: The case of US banks in 2001–2010, *Journal of Applied Econometrics* 31, 1407–1429.

Nguyen, H., L. Simar, and V. Zelenyuk (2022), Data sharpening for improving CLT approximations for DEA-type efficiency estimators, *European Jouranl of Operational Research,* 303, 1469–1480.

Olesen, O. B. and N. C. Petersen (1995), Chance constrained efficiency evaluation, *Management science* 41, 442–457.

Park, B. U., S.-O. Jeong, and L. Simar (2010), Asymptotic distribution of conical-hull estimators of directional edges, *Annals of Statistics* 38, 1320–1340.

Pham, M. D., L. Simar, and V. Zelenyuk (2024), Statistical inference for aggregation of Malmquist productivity indices, *Operations Research* Forthcoming.

Pham, M. D. and V. Zelenyuk (2019), Weak disposability in nonparametric production analysis: A new taxonomy of reference technology sets, *European Journal of Operational Research* 274, 186–198.

Podinovski, V. V. (2022), Variable and constant returns-to-scale production technologies with component processes, *Operations Research* 70, 1238–1258.

Podinovski, V. V. and T. Kuosmanen (2011), Modelling weak disposability in data envelopment analysis under relaxed convexity assumptions, *European Journal of Operational Research* 211, 577–585.

Ray, S. C. (2004), *Data Envelopment Analysis: Theory and Techniques for Economics and Operations Reserach*, Cambridge: Cambridge University Press.

Ruggiero, J. (1996), Efficiency of educational production: An analysis of New York school districts, *The Review of Economics and Statistics* 78, 499–509.

Schmidt, P. and R. C. Sickles (1984), Production frontiers and panel data, *Journal of Business & Economic Statistics* 2, 367–374.

Shephard, R. W. (1953), *Cost and Production Functions*, Princeton: Princeton University Press.

— (1970), *Theory of Cost and Production Functions*, Princeton: Princeton University Press.

Sickles, R. C. and V. Zelenyuk (2019), *Measurement of Productivity and Efficiency*, Cambridge University Press.

Simar, L. and A. Vanhems (2012), Probabilistic characterization of directional distances and their robust versions, *Journal of Econometrics* 166, 342–354.

Simar, L., A. Vanhems, and P. W. Wilson (2012), Statistical inference for DEA estimators of directional distances, *European Journal of Operational Research* 220, 853–864.

Simar, L. and P. W. Wilson (2015), Statistical approaches for non-parametric frontier models: A guided tour, *International Statistical Review* 83, 77–110.

— (2019), Central limit theorems and inference for sources of productivity change measured by nonparametric Malmquist indices, *European Journal of Operational Research* 277, 756–769.

Simar, L. and V. Zelenyuk (2007), Statistical inference for aggregates of Farrell-type efficiencies, *Journal of Applied Econometrics* 27, 1367–1394.

— (2018), Central limit theorems for aggregate efficiency, *Operations Research* 66, 137–149.

— (2020), Improving finite sample approximation by central limit theorems for estimates from data envelopment analysis, *European Journal of Operational Research* 284, 1002–1015.

Simar, L., V. Zelenyuk, and S. Zhao (2023a), Further improvements of finite sample approximation of central limit theorems for envelopment estimators, *Journal of Productivity Analysis* 59, 189–194.

— (2023b), Statistical inference for Hicks–Moorsteen productivity indices. Working Paper Series No. WP08/2023, School of Economics and Centre for Efficiency and Productivity Analysis (CEPA) at The University of Queensland, Australia.

— (2024), Inference for aggregate efficiency: Theory and guidelines for practitioners, *European Journal of Operational Research,* forthcoming.

Tsionas, M., C. F. Parmeter, and V. Zelenyuk (2023), Bayesian artificial neural networks for frontier efficiency analysis, *Journal of Econometrics* 236, 105491.

Wheelock, D. C. and P. W. Wilson (2018), The evolution of scale economies in U.S. banking, *Journal of Applied Econometrics* 33, 16–28.

Wilson, P. W. (2018), Dimension reduction in nonparametric models of production, *European Journal of Operational Research* 267, 349–367.

Zelenyuk, V. and V. Panchenko (2023), Generalized theory for measuring efficiency of individuals and groups, *Annals of Operations Research,* forthcoming.

Zelenyuk, V. and S. Zhao (2023), Further improvements of finite sample approximation of central limit theorems for weighted and unweighted Malmquist productivity indices. Working Paper Series No. WP04/2023, School of Economics and Centre for Efficiency and Productivity Analysis (CEPA) at The University of Queensland, Australia.

Central Limit Theorems for Directional Distance Functions with and without Undesirable Outputs (Online Supplement)

Léopold Simar* Valentin Zelenyuk† Shirong Zhao‡

March 4, 2024

*Simar: Institut de Statistique, Biostatistique et Sciences Actuarielles (ISBA), LIDAM, Université Catholique de Louvain, Voie du Roman Pays 20, B1348 Louvain-la-Neuve, Belgium; email: leopold.simar@uclouvain.be.

†Zelenyuk: School of Economics and Centre for Efficiency and Productivity Analysis (CEPA), University of Queensland, Colin Clark Building (39), St Lucia, Brisbane, Qld 4072, Australia; email: v.zelenyuk@uq.edu.au.

‡Zhao: School of Finance, Dongbei University of Finance and Economics, Dalian, Liaoning 116025, China; email: shironz@163.com.

Contents

List of Tables

Appendix A The Procedures of Estimating the Bias

The steps outlined by Simar and Zelenyuk (2018) to estimate the bias term of $\widehat{\Delta}_n$ are detailed as follows. For $b = 1$, split $\mathcal{X}_n$ into two mutually exclusive and evenly sized subsamples $\mathcal{X}^{(1)}_{n/2}$ and $\mathcal{X}^{(2)}_{n/2}$ (for simplicity, assuming n is even). Then we estimate the simple mean of H_{1i} using the observations only in $\mathcal{X}^{(j)}_{n/2}$, where $j = 1, 2$, and obtain

$$\widehat{\mu}^{(j)}_{1,n,b} = 2n^{-1} \sum_{\{i|(X_i,Y_i)\in\mathcal{X}^{(j)}_{n/2}\}} \widehat{H}_{1i} = 2n^{-1} \sum_{\{i|(X_i,Y_i)\in\mathcal{X}^{(j)}_{n/2}\}} \widehat{\delta}(X_i, Y_i \mid d^i_x, d^i_y, \mathcal{X}^{(j)}_{n/2})H_{2i}. \tag{A.1}$$

Then compute

$$\widehat{\mu}_{1,n,b} = \frac{1}{2}(\widehat{\mu}^{(1)}_{1,n,b} + \widehat{\mu}^{(2)}_{1,n,b}). \tag{A.2}$$

Increasing b by 1 and iterating through the described process for a total of B repetitions, will yield $\{\widehat{\mu}_{1,n,b}\}^B_{b=1}$. The bias term for $\widehat{\mu}_{1,n}$ can be estimated as

$$\widehat{\mathcal{B}}_{\mu_{1,n}} = \sum_{b=1}^{B}(2^\kappa - 1)^{-1}(\widehat{\mu}_{1,n,b} - \widehat{\mu}_{1,n}), \tag{A.3}$$

which leads to the estimate of the bias term for $\widehat{\Delta}_n$, expressed as

$$\widehat{\mathcal{B}}_{\Delta,n} = \frac{\widehat{\mathcal{B}}_{\mu_{1,n}}}{\widehat{\mu}_{2,n}}. \tag{A.4}$$

Appendix B Simulation Results for the Second Scenario

For the second scenorio, we consider the following technology

$$g_i^{\partial} = c \prod_{j=1}^{p} (x_{ij}^{\partial})^{\beta_j}, \tag{B.1}$$

where c is a constant term, x_{ij}^{∂} is the jth efficient input, and g_i^{∂} is the efficient desirable output. The production process for the undesirable output b_i^{∂} is specified as,

$$b_i^{\partial} = 0.2 x_{i1}^{\partial}, \tag{B.2}$$

where for simplicity we assume that the undesirable output (e.g., pollution) is produced using just the first input (e.g., coal). In addition, the ith true directional distance score is generated from $\delta_i \overset{\text{iid}}{\sim} \text{Beta}(1,2)$, i.e., we generate δ_i from the Beta distribution to ensure $\delta_i \in [0,1]$. We first generate the ith observed input as $x_i \overset{\text{iid}}{\sim} \text{Unif}(0,1)$, then calculate the points on the frontier as $(x_i^{\partial}, g_i^{\partial}, b_i^{\partial})$, where $x_i^{\partial} = x_i(1-\delta_i)$, $g_i^{\partial} = c\prod_{j=1}^{p}(x_{ij}^{\partial})^{\beta_j}$, and $b_i^{\partial} = 0.2x_{i1}^{\partial}$. Finally, we set the ith observed desirable output as $g_i = g_i^{\partial}/(1+\delta_i)$ and observed undesirable output as $b_i = b_i^{\partial}/(1-\delta)$. Thus, a simulated sample $\mathcal{X}_n = \{(x_i, g_i, b_i)\}_{i=1}^{n}$ is created. The values of the other parameters (e.g., c and β_js) are the same as the first scenario in the main text.

Table B.1 shows the simulation results, which are similar to those in the main text for the first scenario.

Table B.1: Coverages of Estimated Confidence Intervals for Aggregate Directional Distance Functions using VRS-DEA Estimator for the Second Scenario

Panel A: The Cases with $p + q_g + q_r < 4$

				0.90		0.95		0.99	
p	q_g	q_b	n	(i)	(ii)	(i)	(ii)	(i)	(ii)
1	1	1	10	0.439	0.623	0.525	0.703	0.669	0.805
1	1	1	20	0.535	0.732	0.615	0.799	0.782	0.882
1	1	1	50	0.633	0.762	0.736	0.849	0.870	0.940
1	1	1	100	0.698	0.798	0.802	0.873	0.916	0.960
1	1	1	200	0.710	0.837	0.806	0.917	0.931	0.966
1	1	1	500	0.761	0.848	0.852	0.917	0.950	0.983
1	1	1	1000	0.790	0.893	0.882	0.930	0.967	0.985

Panel B: The Cases with $p + q_g + q_r \geq 4$

				0.90			0.95			0.99		
p	q_g	q_b	n	(i)	(ii)	(iii)	(i)	(ii)	(iii)	(i)	(ii)	(iii)
2	1	1	10	0.132	0.460	0.549	0.168	0.527	0.624	0.309	0.643	0.719
2	1	1	20	0.108	0.599	0.697	0.167	0.670	0.783	0.324	0.806	0.889
2	1	1	50	0.092	0.709	0.802	0.153	0.790	0.883	0.310	0.894	0.956
2	1	1	100	0.111	0.722	0.808	0.169	0.795	0.880	0.369	0.905	0.959
2	1	1	200	0.115	0.730	0.812	0.173	0.823	0.883	0.380	0.914	0.967
2	1	1	500	0.118	0.744	0.841	0.177	0.820	0.907	0.376	0.938	0.974
2	1	1	1000	0.118	0.742	0.858	0.185	0.830	0.909	0.395	0.936	0.979
3	1	1	10	0.021	0.303	0.403	0.036	0.342	0.463	0.092	0.423	0.552
3	1	1	20	0.010	0.454	0.623	0.022	0.525	0.701	0.070	0.649	0.810
3	1	1	50	0.004	0.601	0.759	0.009	0.679	0.839	0.027	0.807	0.930
3	1	1	100	0.000	0.629	0.829	0.000	0.713	0.890	0.011	0.838	0.966
3	1	1	200	0.000	0.586	0.817	0.000	0.675	0.881	0.008	0.818	0.968
3	1	1	500	0.000	0.525	0.842	0.000	0.636	0.908	0.000	0.792	0.976
3	1	1	1000	0.000	0.455	0.835	0.000	0.547	0.903	0.001	0.736	0.974
4	1	1	10	0.004	0.147	0.248	0.008	0.179	0.285	0.017	0.223	0.363
4	1	1	20	0.000	0.273	0.491	0.000	0.317	0.547	0.005	0.413	0.657
4	1	1	50	0.000	0.450	0.708	0.000	0.527	0.809	0.001	0.649	0.916
4	1	1	100	0.000	0.526	0.819	0.000	0.614	0.894	0.000	0.751	0.966
4	1	1	200	0.000	0.521	0.838	0.000	0.603	0.914	0.000	0.747	0.979
4	1	1	500	0.000	0.418	0.839	0.000	0.488	0.912	0.000	0.630	0.970
4	1	1	1000	0.000	0.252	0.832	0.000	0.321	0.904	0.000	0.456	0.970

Appendix C More Details on the Dimension Reduction Used in the Empirical Illustration

In this appendix, we present some details not shown in the paper for the empirical illustration.

Following Malikov et al. (2016) and Wheelock and Wilson (2018), five desirable output quantities are specified for the banks' production function: g_1 = consumer loans, g_2 = real estate loans, g_3 = commercial and industrial loans, g_4 = securities, and g_5 = off-balance-sheet items. Further, five input quantities are specified: x_1 = purchased funds, x_2 = physical capital, x_3 = the number of full-time employees, x_4 = interest-bearing transaction accounts, and x_5 = non-transaction accounts. Moreover, we specify one undesirable output: b = the non-performing loans. Altogether, we have a total of $2,397$ bank-year observations with the number of banks in each year ranging from 179 to 277 over 2001–2010. The estimation results are obtained by estimating the technology frontiers for each year in the sample.

Before presenting the results, we first address the necessity to reduce the dimensions of the data. Theorem 4 shows that given the sample size, the estimation errors for the non-parametric efficiency estimators increase with the dimensions. In our specification, we have 11 dimensions in total while the sample size in each year is no larger than 277. To enhance the accuracy of the estimation, we first consolidate three types of loans (g_1, g_2, and g_3) into a single aggregated loan (g_6), and similarly, we combine two types of accounts (x_4 and x_5) into a unified account (x_6); by implementing this approach, we form the matrix for the desirable outputs $\boldsymbol{G} = [g_4, g_5, g_6]$ and the matrix for the inputs $\boldsymbol{X} = [x_1, x_2, x_3, x_6]$. Subsequently, we employ the dimension reduction method introduced by Daraio and Simar (2007) to reduce the dimensions for both $\boldsymbol{G}$ and $\boldsymbol{X}$, separately for each year. Our computations show that the ratio of the largest eigenvalue to the sum of the eigenvalues for the matrix $\boldsymbol{G}'\boldsymbol{G}$ is no less than 0.9004 for each year; and it is no less than 0.9437 in each year for the matrix $\boldsymbol{X}'\boldsymbol{X}$. These two values fall comfortably within the guidelines outlined in Wilson (2018) and hence we use only the first principal components for the matrices of the inputs and desirable outputs, which are denoted as x^* and g^*, respectively. Through this process, we obtain the aggregated input-output data, $\mathcal{X}_n = \{(x_i^*, g_i^*, b_i)\}_{i=1}^n$, where $p = 1$, $q_g = 1$ and $q_b = 1$.

References

Daraio, C. and L. Simar (2007), *Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications*, New York: Springer Science & Business Media.

Malikov, E., S. C. Kumbhakar, and M. G. Tsionas (2016), A cost system approach to the stochastic directional technology distance function with undesirable outputs: The case of US banks in 2001–2010, *Journal of Applied Econometrics* 31, 1407–1429.

Simar, L. and V. Zelenyuk (2018), Central limit theorems for aggregate efficiency, *Operations Research* 66, 137–149.

Wheelock, D. C. and P. W. Wilson (2018), The evolution of scale economies in U.S. banking, *Journal of Applied Econometrics* 33, 16–28.

Wilson, P. W. (2018), Dimension reduction in nonparametric models of production, *European Journal of Operational Research* 267, 349–367.