STATISTICAL INFERENCE FOR AGGREGATION OF MALMQUIST PRODUCTIVITY INDICES

Manh Pham, Léopold Simar, Valentin Zelenyuk

REPRINT | 2023 / 10

UCLouvain

LIDAM
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

informs
https://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Articles in Advance, pp. 1–15
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Methods

Statistical Inference for Aggregation of Malmquist Productivity Indices

Manh Pham,[a,*] Léopold Simar,[b] Valentin Zelenyuk[c]

[a] School of Economics, The University of Queensland, Brisbane QLD 4072, Australia; [b] Institut de Statistique, Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, B1348 Louvain-la-Neuve, Belgium; [c] School of Economics and Centre for Efficiency and Productivity Analysis, The University of Queensland, Brisbane QLD 4072, Australia

*Corresponding author

Contact: ducmanh.pham@uq.net.au, https://orcid.org/0000-0001-8970-9443 (MP); leopold.simar@uclouvain.be, https://orcid.org/0000-0003-0791-8490 (LS); v.zelenyuk@uq.edu.au, https://orcid.org/0000-0001-6120-4291 (VZ)

Received: July 14, 2019
Revised: February 3, 2022; July 18, 2022
Accepted: November 19, 2022
Published Online in Articles in Advance: January 13, 2023

Area of Review: Optimization

https://doi.org/10.1287/opre.2022.2424

Abstract. The Malmquist productivity index (MPI) has gained popularity among studies on the dynamic change of productivity of decision-making units (DMUs). In practice, this index is frequently reported at aggregate levels (e.g., public and private firms) in the form of simple, equally weighted arithmetic or geometric means of individual MPIs. A number of studies emphasize that it is necessary to account for the relative importance of individual DMUs in the aggregations of indices in general and of the MPI in particular. Whereas more suitable aggregations of MPIs have been introduced in the literature, their statistical properties have not been revealed yet, preventing applied researchers from making essential statistical inferences, such as confidence intervals and hypothesis testing. In this paper, we fill this gap by developing a full asymptotic theory for an appealing aggregation of MPIs. On the basis of this, meaningful statistical inferences are proposed, their finite-sample performances are verified via extensive Monte Carlo experiments, and the importance of the proposed theoretical developments is illustrated with an empirical application to real data.

Funding: M. Pham acknowledges support from an Australian Government Research Training Program Scholarship. V. Zelenyuk acknowledges financial support from the University of Queensland and the Australian Research Council [Grant FT170100401].
Supplemental Material: The e-companion is available at https://doi.org/10.1287/opre.2022.2424.

Keywords: aggregation • asymptotics • DEA • hypothesis test • inference • Malmquist index • productivity

1. Introduction

The Malmquist productivity index (MPI), since its introduction by Caves et al. (1982), has become one of the most widely used tools for analyzing the performance of decision-making units (DMUs) in terms of their productivity change over time. A large number of applications of the MPI has been deployed in a wide variety of areas, including agriculture, airlines, banking, electric utilities, healthcare, insurance companies, and public sectors, to mention just a few.[1] Among the approaches to estimate MPI, data envelopment analysis (DEA) appears to be the most popular in the literature.[2]

In practice, apart from measuring the productivity change of a single DMU, there is also an essential need for analyzing productivity change at an aggregate level (e.g., firms grouped by ownership status, such as public and private). For instance, when studying performances of European and U.S. banking systems, Pastor et al. (1997) use the median, simple average, and weighted (by total assets) average of individual MPIs of banks in each country to make the cross-country comparison. Another example is Tortosa-Ausina et al. (2008), who examine productivity growth and the productive efficiency of Spanish savings banks over the period 1992–1998 and report both the simple arithmetic and geometric means of the MPIs of individual banks. When it comes to aggregation of indices, it is also pointed out that simple averages (i.e., arithmetic and geometric means), which assign the same weight to each individual regardless of its relative economic significance (e.g., market share), might lead to very different conclusions in relation to the averages that account for economic weight (e.g., see Ylvinger 2000, Ebert and Welsch 2004).

Accounting for economic weights of DMUs based on their relative importance is also consistent with reality in which many industries are usually dominated by a few firms. For example, according to the U.S. Federal Reserve, in 2019, about 40% of the total domestic assets of 1,835 large commercial banks in the United States belonged to four large banks: JP Morgan Chase, Bank of America, Wells Fargo, and Citibank.[3] Hence, the aggregate functional forms and weights should be considered carefully in order to achieve meaningful aggregations. This consideration also encompasses productivity and efficiency indices in general and the MPI in particular (e.g., see Färe and Zelenyuk 2003; Färe and Grosskopf 2004; Zelenyuk 2006; ten Raa 2011; Wang et al. 2017; Walheer 2018, 2019).

Zelenyuk (2006) proposes two aggregations of MPIs that resemble the weighted harmonic-type and geometric means of efficiencies, accounting for the economic importance of individuals. Until now, their statistical properties have not been unveiled, preventing applied researchers from making well-grounded statistical inferences, such as confidence intervals and hypothesis testing, which are important in practice. Leveraging Kneip et al. (2015), Simar and Zelenyuk (2018), and Kneip et al. (2021) (hereafter KSW2015, SZ2018, and KSW2021, respectively), we bridge this gap by developing a comprehensive asymptotic theory for the weighted harmonic-type mean aggregation of MPIs in two contexts: (i) individual efficiency scores are observable, and (ii) individual efficiency scores are nonobservable and estimated via DEA relative to the conical hulls of the production technology sets. These new developments enable applied researchers to obtain meaningful statistical inferences on aggregate productivity change measured by MPIs.[4]

It is important to note that the complexity of the used statistic implies that the traditional delta method, on which KSW2015, SZ2018, and KSW2021 are based, leaves the stochastic remainder of order $O_p(n^{-1/2})$, which does not provide sufficiently tight bounds in our context. Interestingly, we find that utilizing the uniform version of the delta method helps in canceling out the $O_p(n^{-1/2})$ term, thus reducing the stochastic remainder to $o_p(n^{-1/2})$. To the best of our knowledge, this approach is a novel one compared with the previous works of KSW2015, SZ2018, and KSW2021 and opens the path for deriving the asymptotic properties of a variety of sophisticated indices, such as the weighted geometric mean aggregation of MPIs and the Hicks–Moorsteen productivity index.

Finally, we also conduct Monte Carlo experiments to verify the performance of the newly developed statistical inferences in finite samples as well as illustrate it on real data.

2. Preliminaries

We denote inputs and outputs by column vectors $x \in \mathbb{R}_+^p$ and $y \in \mathbb{R}_+^q$, respectively. The production technology set at time t is defined as

$$\Psi^t = \{(x,y) \in \mathbb{R}_+^p \times \mathbb{R}_+^q : x \text{ can produce } y \text{ at time } t\},\ t = 1,2, \tag{1}$$

which is assumed to satisfy common regularity assumptions (see Appendix A for details).

To date, a number of efficiency measures have been proposed to evaluate the performance of a particular DMU relative to a production technology set, and the Farrell-type efficiency measures appear to be the most popular in the literature (Farrell 1957). They are also an important component in the decomposition of profit efficiency (Färe et al. 2019). Apart from this, the hyperbolic efficiency measure (Färe et al. 1985) is also appealing as it seeks to decrease input quantities and increase output quantities simultaneously and equiproportionally. Formally, the efficiency of a DMU with an input–output combination $z = (x,y) \in \mathbb{R}_+^p \times \mathbb{R}_+^q$ can be evaluated relative to the production technology set Ψ^t via these measures as follows:

- Farrell-type output-oriented efficiency measure:

$$\lambda(z|\Psi^t) = \lambda(x,y|\Psi^t) = \sup_{\lambda}\{\lambda : (x, \lambda y) \in \Psi^t\}. \tag{2}$$

- Hyperbolic efficiency measure:

$$\gamma(z|\Psi^t) = \gamma(x,y|\Psi^t) = \inf_{\gamma}\{\gamma > 0 : (\gamma x, \gamma^{-1} y) \in \Psi^t\}. \tag{3}$$

By construction, for all $z = (x,y) \in \Psi^t$, we have $\lambda(z|\Psi^t) \geq 1$ and $0 \leq \gamma(z|\Psi^t) \leq 1$. Now, we define the conical hull of the set Ψ^t as[5]

$$\mathcal{C}(\Psi^t) = \{(ax, ay) : (x,y) \in \Psi^t, a \in \mathbb{R}_+^1\}. \tag{4}$$

Obviously, $\Psi^t \subseteq \mathcal{C}(\Psi^t)$. Conventionally, Ψ^t is said to exhibit globally constant returns to scale (CRS) if $\Psi^t = \mathcal{C}(\Psi^t)$ and variable returns to scale (VRS) otherwise (which means Ψ^t might exhibit increasing, constant, or decreasing returns to scale in some local regions).

The conical Farrell-type output-oriented efficiency measure, denoted by λ_C, for a DMU with an input–output combination $z = (x,y) \in \mathbb{R}_+^p \times \mathbb{R}_+^q$ is defined as

$$\begin{aligned}\lambda_C(z|\Psi^t) &= \lambda_C(x,y|\Psi^t) = \lambda(x,y|\mathcal{C}(\Psi^t)) \\ &= \sup_{\lambda}\{\lambda : (x, \lambda y) \in \mathcal{C}(\Psi^t)\},\end{aligned} \tag{5}$$

and γ_C can also be defined in a way similar to λ_C.

Now suppose that the input–output combinations of the interested DMU observed in periods 1 and 2 are $z^1 \in \Psi^1$ and $z^2 \in \Psi^2$, respectively; then, the output-oriented conical MPI, which measures the productivity change of this DMU from period 1 to 2, can be defined as

$$M_O(z^1, z^2) = \left(\frac{\lambda_C(z^2|\Psi^1)}{\lambda_C(z^1|\Psi^1)} \times \frac{\lambda_C(z^2|\Psi^2)}{\lambda_C(z^1|\Psi^2)}\right)^{-1/2}. \tag{6}$$

It is clear that $M_O(z^1, z^2)$ take values in $(0, \infty)$. In particular, values in $(1, \infty)$, $\{1\}$, and (0, 1) indicate that the productivity of firm i has improved, remained constant, and deteriorated from period 1 to 2, respectively.[6]

Similar to KSW2021, here we emphasize the importance of measuring efficiency relative to the conical hulls of the production technology sets rather than the sets themselves. On the one hand, Grifell-Tatjé and Lovell

(1995) indicate that, under VRS, the MPI does not account for productivity change accurately because of a systematic bias. In addition, Ray and Desli (1997), when analyzing productivity changes of 17 Organisation for Economic Co-operation and Development countries, noted that computing MPI for some countries under VRS might be infeasible. On the other hand, it might be too restrictive to impose CRS on the production technology as discussed in, for example, KSW2021. Interestingly, using the conical hull of the production technology set can solve this dilemma because measuring MPI relative to the conical hull is always feasible, and furthermore, this approach allows the true production technology set to exhibit VRS rather than the more restricted CRS. Therefore, we focus on the conical Farrell-type output-oriented MPI in this paper, noting that similar results for the other orientations (e.g., input-oriented) can be developed analogously.

Now, consider a sample $\mathcal{X}_n = \mathcal{X}_n^1 \cup \mathcal{X}_n^2$ consisting of n DMUs observed in period 1 (i.e., $\mathcal{X}_n^1$) and period 2 (i.e., $\mathcal{X}_n^2$). More precisely, for each $t = 1, 2$, $\mathcal{X}_n^t = \{Z_i^t\}_{i=1}^n$, where $Z_i^t = (X_i^t, Y_i^t)$, X_i^t and Y_i^t are column vectors of inputs and outputs of DMU i $(i = 1, \ldots, n)$, respectively.

A common approach to aggregate individual MPIs is to use the equally weighted arithmetic or geometric mean, in which the latter seems to dominate the former because of the multiplicative essence of the MPI. These types of aggregations treat individual DMUs equally and, hence, ignore their relative economic importance (e.g., market share), which motivates us to investigate the aggregate MPIs.

We assume that all DMUs in the same time period face common output prices (i.e., "law of one price"). The statistical theory developed here still applies to the context in which DMUs face different prices, and the assumption of the law of one price (or common equilibrium) is required to maintain the Koopmans-type theorem of aggregation upon which the theory of aggregate efficiency and productivity are built.

Here, we develop statistical theory for the following aggregate MPI:

$$\overline{M} = \left(\frac{\sum_{i=1}^n \beta_i^2 \lambda_C(Z_i^2 | \Psi^1)}{\sum_{i=1}^n \beta_i^1 \lambda_C(Z_i^1 | \Psi^1)} \times \frac{\sum_{i=1}^n \beta_i^2 \lambda_C(Z_i^2 | \Psi^2)}{\sum_{i=1}^n \beta_i^1 \lambda_C(Z_i^1 | \Psi^2)} \right)^{-1/2}, \tag{7}$$

where $\beta_i^t = \frac{w^t Y_i^t}{w^t \sum_{i=1}^n Y_i^t}$ $(i = 1, \ldots, n)$ are economic weights and $w^t \in \mathbb{R}_{++}^q$ are the row vector of output prices in the period t $(t = 1, 2)$.[7]

3. DEA Estimators from the Statistical Viewpoint

The technology sets Ψ^1 and Ψ^2 as well as efficiency measures and indices are unobserved in reality and must be estimated from data, raising the need for respective statistical inferences. Among estimation methodologies to date, DEA has emerged as one of the most popular, attracting numerous theoretical and empirical attention.[8] More specifically, DEA appears to be the most popular method to estimate MPI, especially since the seminal work of Färe et al. (1994). Prior to presenting the estimation details and setting up a statistical model for the DEA-based MPI, we need additional Assumptions A.5–A.10 in Appendix A. These assumptions are output-oriented analogues of the input-oriented assumptions 2.4–2.7, 3.1, and 3.2 in KSW2021.

By virtue of these assumptions, the sample $\mathcal{X}_n$ can be viewed as a random set generated from Ψ^1 and Ψ^2. For $t = 1, 2$, Ψ^t and $\mathcal{C}(\Psi^t)$ can be estimated via DEA-type estimators as follows:[9]

$$\widehat{\Psi}_n^t = \left\{ (x, y) \in \mathbb{R}_+^p \times \mathbb{R}_+^q : x \geqq \sum_{i=1}^n X_i^t \zeta_i, y \leqq \sum_{i=1}^n Y_i^t \zeta_i, \sum_{i=1}^n \zeta_i = 1, \zeta_i \in \mathbb{R}_+^1, \ \forall i = 1, \ldots, n \right\}, \tag{8}$$

$$\mathcal{C}(\widehat{\Psi}_n^t) = \left\{ (x, y) \in \mathbb{R}_+^p \times \mathbb{R}_+^q : x \geqq \sum_{i=1}^n X_i^t \zeta_i, y \leqq \sum_{i=1}^n Y_i^t \zeta_i, \zeta_i \in \mathbb{R}_+^1, \ \forall i = 1, \ldots, n \right\}. \tag{9}$$

Substituting Ψ^t by $\widehat{\Psi}_n^t$ in (2) and (5) gives the DEA-type estimators of the Farrell-type output-oriented efficiency measure and its conical hull version:

$$\lambda(z | \widehat{\Psi}_n^t) = \lambda(x, y | \widehat{\Psi}_n^t) = \sup_\lambda \left\{ \lambda : x \geqq \sum_{i=1}^n X_i^t \zeta_i, \lambda y \leqq \sum_{i=1}^n Y_i^t \zeta_i, \sum_{i=1}^n \zeta_i = 1, \zeta_i \in \mathbb{R}_+^1, \ \forall i = 1, \ldots, n \right\},$$

$$\lambda_C(z | \widehat{\Psi}_n^t) = \lambda_C(x, y | \widehat{\Psi}_n^t) = \sup_\lambda \left\{ \lambda : x \geqq \sum_{i=1}^n X_i^t \zeta_i, \lambda y \leqq \sum_{i=1}^n Y_i^t \zeta_i, \zeta_i \in \mathbb{R}_+^1, \ \forall i = 1, \ldots, n \right\}, \tag{10}$$

respectively. Note that, because $\widehat{\Psi}_n^t$ is, in turn, determined by $\mathcal{X}_n^t$, henceforward in this paper we use the notations $\lambda(z | \mathcal{X}_n^t)$ and $\lambda_C(z | \mathcal{X}_n^t)$ instead of $\lambda(z | \widehat{\Psi}_n^t)$ and $\lambda_C(z | \widehat{\Psi}_n^t)$, respectively, to emphasize the data set from which the estimators are computed.

The aforementioned aggregate MPI can be estimated via DEA as follows:

$$\widehat{\overline{M}}(\mathcal{X}_n) = \left(\frac{\sum_{i=1}^n \beta_i^2 \lambda_C(Z_i^2 | \mathcal{X}_n^1)}{\sum_{i=1}^n \beta_i^1 \lambda_C(Z_i^1 | \mathcal{X}_n^1)} \times \frac{\sum_{i=1}^n \beta_i^2 \lambda_C(Z_i^2 | \mathcal{X}_n^2)}{\sum_{i=1}^n \beta_i^1 \lambda_C(Z_i^1 | \mathcal{X}_n^2)} \right)^{-1/2}. \tag{11}$$

KSW2015 are the first to discover the asymptotic properties of the Farrell-type technical efficiency evaluated at random points. In addition, they propose important central limit theorems for the arithmetic mean of the efficiency of random samples, enabling researchers to make inference about the efficiency of groups of firms. Most recently, KSW2021 extend the work of KSW2015 to the dynamic context and provide the following result.

Lemma 1. *Under Assumptions A.1–A.10, as $n \to \infty$,*

$$E\left(\log\gamma_C(Z_i^s|\mathcal{X}_n^t)-\log\gamma_C(Z_i^s|\Psi^t)\right)=\widetilde{C}_{st}n^{-\kappa}+R_{n,\kappa}, \quad (12)$$

$$E\left([\log\gamma_C(Z_i^s|\mathcal{X}_n^t)-\log\gamma_C(Z_i^s|\Psi^t)]^2\right)=o(n^{-\kappa}), \quad (13)$$

$$|E([\log\gamma_C(Z_i^s|\mathcal{X}_n^t)-E(\log\gamma_C(Z_i^s|\mathcal{X}_n^t))]\times \times[\log\gamma_C(Z_j^{s^*}|\mathcal{X}_n^{t^*})-E(\log\gamma_C(Z_j^{s^*}|\mathcal{X}_n^{t^*}))])|=o(n^{-1}), \quad (14)$$

for all $i,j\in\{1,\dots,n\}, i\neq j$; $s,t,s^*,t^*\in\{1,2\}$, where $\widetilde{C}_{st}$ is a constant, $R_{n,\kappa}$ is a remainder of order smaller than $n^{-\kappa}$, and $\kappa=2/(p+q+1)$ if the true technology Ψ^t exhibits VRS and $\kappa=2/(p+q)$ if the true technology Ψ^t exhibits CRS (i.e., if $\Psi^t=\mathcal{C}(\Psi^t)$).

These results pave the way for statistical inference in a dynamic context. Particularly, KSW2021 derive asymptotic properties for DEA-based estimators of MPIs for individual DMUs as well as their geometric mean. Our goal is to generalize their results to the case in which researchers want to account for economic weights in the aggregation.

4. Asymptotic Properties When the True Efficiency Is Unknown

Consider the log version of $\overline{M}$ given by

$$\log\overline{M}=-\frac{1}{2}\left[\log\left(\sum_{i=1}^n\beta_i^2\lambda_C(Z_i^2|\Psi^1)\right)+\log\left(\sum_{i=1}^n\beta_i^2\lambda_C(Z_i^2|\Psi^2)\right)-\log\left(\sum_{i=1}^n\beta_i^1\lambda_C(Z_i^1|\Psi^1)\right)-\log\left(\sum_{i=1}^n\beta_i^1\lambda_C(Z_i^1|\Psi^2)\right)\right], \quad (15)$$

and to make our notation more concise, for $i=1,\dots,n$, let

$$U_{1,i}=\lambda_C(Z_i^2|\Psi^1)w^2Y_i^2,\ U_{2,i}=\lambda_C(Z_i^2|\Psi^2)w^2Y_i^2,\ U_{3,i}=\lambda_C(Z_i^1|\Psi^1)w^1Y_i^1,\ U_{4,i}=\lambda_C(Z_i^1|\Psi^2)w^1Y_i^1,\ U_{5,i}=w^2Y_i^2,\ U_{6,i}=w^1Y_i^1. \quad (16)$$

Clearly, $U_{s,i}$ are scalar-valued random variables for all $s=1,2,\dots,6$ and $i=1,\dots,n$. Denote $\mu_s=E(U_{s,i})$ and $\widehat{\mu}_{s,n}=n^{-1}\sum_{i=1}^n U_{s,i}$ $(s=1,\dots,6)$. Similar to SZ2018, we have

$$\sum_{i=1}^n\beta_i^2\lambda_C(Z_i^2|\Psi^1)=\sum_{i=1}^n\frac{w^2Y_i^2}{w^2\sum_{i=1}^nY_i^2}\lambda_C(Z_i^2|\Psi^1)=\frac{\sum_{i=1}^n\lambda_C(Z_i^2|\Psi^1)w^2Y_i^2}{\sum_{i=1}^nw^2Y_i^2}=\frac{\sum_{i=1}^nU_{1,i}}{\sum_{i=1}^nU_{5,i}}=\frac{\widehat{\mu}_{1,n}}{\widehat{\mu}_{5,n}}.$$

Analogously,

$$\sum_{i=1}^n\beta_i^2\lambda_C(Z_i^2|\Psi^2)=\frac{\widehat{\mu}_{2,n}}{\widehat{\mu}_{5,n}},\quad \sum_{i=1}^n\beta_i^1\lambda_C(Z_i^1|\Psi^1)=\frac{\widehat{\mu}_{3,n}}{\widehat{\mu}_{6,n}},\quad \sum_{i=1}^n\beta_i^1\lambda_C(Z_i^1|\Psi^2)=\frac{\widehat{\mu}_{4,n}}{\widehat{\mu}_{6,n}}.$$

Consequently, the aforementioned aggregation of MPIs can be expressed concisely as

$$\log\overline{M}=\frac{1}{2}(\log\widehat{\mu}_{3,n}+\log\widehat{\mu}_{4,n}-\log\widehat{\mu}_{1,n}-\log\widehat{\mu}_{2,n})+\log\widehat{\mu}_{5,n}-\log\widehat{\mu}_{6,n}. \quad (17)$$

Therefore, $\log\overline{M}$ is a point estimate of the following parameter:

$$\xi=\frac{1}{2}(\log\mu_3+\log\mu_4-\log\mu_1-\log\mu_2)+\log\mu_5-\log\mu_6, \quad (18)$$

whereas $\overline{M}$ is a point estimate of $\exp(\xi)$. Thus, hereafter, we write $\log\overline{M}$ as $\widehat{\xi}_n$ and develop an asymptotic theory for the statistical inference of the parameters of interest ξ.

A key stepping stone for developing the asymptotic properties of DEA-estimated aggregate MPIs is to first develop such results for an "ideal scenario" in which the true efficiency scores ($\lambda_C(\cdot|\Psi^1)$ and $\lambda_C(\cdot|\Psi^2)$), which enter the aggregate MPIs in various ways, are known. In principle, this task is a standard application of the delta method although somewhat tedious because of the complexity of the aggregate MPI formulae (e.g., relative to the aggregate efficiency in SZ2018). To the best of our knowledge, this step has not been accomplished before, and so this work is the first in the literature to do so. To save space, we present these developments of asymptotic theory in the ideal scenario in Appendix B. Meanwhile, in this section, we take those developments further by relaxing the assumption of knowledge of $\lambda_C(\cdot|\Psi^1)$ and $\lambda_C(\cdot|\Psi^2)$ and derive new central limit theorems that allow researchers to make statistical inferences about aggregate MPIs based on feasible DEA-type estimators. To do so, by analogy with the preceding definitions, for $i=1,\dots,n$, let

$$\widehat{U}_{1,i}=\lambda_C(Z_i^2|\mathcal{X}_n^1)w^2Y_i^2,\ \widehat{U}_{2,i}=\lambda_C(Z_i^2|\mathcal{X}_n^2)w^2Y_i^2,$$

$$\widehat{U}_{3,i}=\lambda_C(Z_i^1|\mathcal{X}_n^1)w^1Y_i^1,\ \widehat{U}_{4,i}=\lambda_C(Z_i^1|\mathcal{X}_n^2)w^1Y_i^1,$$

where we recall that $\lambda_C(.|.)$ is defined before in (10). Next, let

$$\widehat{\widehat{\mu}}_{s,n} = n^{-1}\sum_{i=1}^{n} \widehat{U}_{s,i}, \quad s = 1,\ldots,4. \tag{19}$$

Because $\widehat{\widehat{\mu}}_{s,n}$ $(s = 1,\ldots,4)$, $\widehat{\mu}_{5,n}$ and $\widehat{\mu}_{6,n}$ can be computed from the data, the estimators

$$\begin{aligned}\widehat{\widehat{\xi}}_n &= \log\left(\widehat{\widehat{M}}(\mathcal{X}_n)\right)\\ &= \frac{1}{2}(\log\widehat{\widehat{\mu}}_{3,n} + \log\widehat{\widehat{\mu}}_{4,n} - \log\widehat{\widehat{\mu}}_{1,n} - \log\widehat{\widehat{\mu}}_{2,n}) + \log\widehat{\mu}_{5,n}\\ &\quad - \log\widehat{\mu}_{6,n}\end{aligned}$$

are feasible. Hence, we develop an asymptotic theory for $\widehat{\widehat{\xi}}_n$ in order to make feasible statistical inferences for the corresponding parameters of interest ξ and $\exp(\xi)$.

To facilitate further discussion on asymptotic rates, let $v_{n,\kappa}$ be a sequence of positive real numbers $\{v_{n,\kappa}\}_{n=1}^{\infty}$ defined by

$$v_{n,\kappa} = \left(\frac{\log n}{n}\right)^{\frac{3}{p+q+d}}, \tag{20}$$

where, in general, κ corresponds to the type of the estimator deployed. Specifically, recall that, for DEA, $\kappa = 2/(p+q+d)$ with $d=1$ when the true technology Ψ^t exhibits VRS and $d=0$ when the true technology Ψ^t exhibits CRS (i.e., if $\Psi^t = \mathcal{C}(\Psi^t)$). For the sake of brevity, in this paper, we focus on $d=1$ (as in KSW2021).

Theorem 1 establishes the basic asymptotic properties of moments of $\widehat{U}_{s,i}$ for $s = 1,\ldots,4$ and $i = 1,\ldots,n$, and its proof can be found in E-companion EC.1.[10]

Theorem 1. *Under Assumptions A.1–A.10, there exist constants $C_s \in (0,\infty)$ such that, as $n \to \infty$,*

$$E(\widehat{U}_{s,i} - U_{s,i}) = C_s n^{-\kappa} + O(v_{n,\kappa}), \tag{21}$$

$$E([\widehat{U}_{s,i} - U_{s,i}]^2) = o(n^{-\kappa}), \tag{22}$$

$$|E([\widehat{U}_{s,i} - E(\widehat{U}_{s,i})][\widehat{U}_{t,j} - E(\widehat{U}_{t,j})])| = o(n^{-1}), \tag{23}$$

for all $i,j \in \{1,\ldots,n\}, i \neq j; s,t \in \{1,\ldots,4\}$.

The following theorem develops Theorem 1 and provides essential tools for deriving the asymptotic properties of $\widehat{\mu}_{s,n}$ $(s = 1,\ldots,4)$ in later stages.

Theorem 2. *Under Assumptions A.1–A.10, as $n \to \infty$,*

i. $E(\widehat{U}_{s,i}) = \mu_s + C_s n^{-\kappa} + O(v_{n,\kappa})$, (24)

ii. $Cov(\widehat{U}_{t,i}, \widehat{U}_{t^*,i}) = \sigma_{tt^*} + o(n^{-\kappa/2})$, (25)

iii. $Cov(\widehat{U}_{s,i}, U_{r,i}) = \sigma_{sr} + o(n^{-\kappa/2})$, (26)

for all $i \in \{1,\ldots,n\}$, $s,t,t^* \in \{1,\ldots,4\}$, $r \in \{5,6\}$; C_s are the same constants as in Theorem 1.

A proof of this theorem is relatively long and so is deferred to E-companion EC.1, whereas here, it is important to note that this theorem is more comprehensive than lemma 1 of SZ2018 because it encompasses the asymptotic covariance of the two estimators containing efficiency scores (i.e., $Cov(\widehat{U}_{t,i}, \widehat{U}_{t^*,i})$ in Theorem 2(ii)), which we solve by decomposing the covariance into four components and examining the asymptotic behavior of each one separately.

To simplify further our notation, let $\widetilde{\mu}_{s,n} = E(\widehat{\widehat{\mu}}_{s,n})$ for $s = 1,\ldots,4$. In the next theorem, we establish important properties of $\widehat{\widehat{\mu}}_{s,n}$ and consistency of estimators of σ_{st}.[11]

Theorem 3. *Under Assumptions A.1–A.10, as $n \to \infty$,*

i. $\widetilde{\mu}_{s,n} = \mu_s + C_s n^{-\kappa} + O(v_{n,\kappa})$, (27)

ii. $\widehat{\widehat{\mu}}_{s,n} - \widetilde{\mu}_{s,n} = \widehat{\mu}_{s,n} - \mu_s + o_p(n^{-1/2})$, (28)

iii. $\sqrt{n}(\widehat{\widehat{\mu}}_{s,n} - \widetilde{\mu}_{s,n}) \xrightarrow{d} \mathcal{N}(0, \sigma_{ss})$, (29)

iv. $\widehat{\widehat{\sigma}}_{st,n} = n^{-1}\sum_{i=1}^{n}(\widehat{U}_{s,i} - \widehat{\widehat{\mu}}_{s,n})(\widehat{U}_{t,i} - \widehat{\widehat{\mu}}_{t,n}) \xrightarrow{p} \sigma_{st}$, (30)

v. $\widehat{\widehat{\sigma}}_{sr,n} = n^{-1}\sum_{i=1}^{n}(\widehat{U}_{s,i} - \widehat{\widehat{\mu}}_{s,n})(U_{r,i} - \widehat{\mu}_{r,n}) \xrightarrow{p} \sigma_{sr}$, (31)

vi. $\widehat{\sigma}_{rr^*,n} = n^{-1}\sum_{i=1}^{n}(U_{r,i} - \widehat{\mu}_{r,n})(U_{r^*,i} - \widehat{\mu}_{r^*,n}) \xrightarrow{p} \sigma_{rr^*}$ (32)

for $s,t \in \{1,\ldots,4\}$, $r,r^* \in \{5,6\}$.

A proof of this theorem is also relatively long and so is deferred to E-companion EC.1. Meanwhile, it is important to note that, compared with theorem 1 of SZ2018, Theorem 3 includes an additional result about the covariance of random variables containing the efficiency scores (i.e., part (iv)). Theorem 3 helps us derive a consistent estimator of V_ξ. Indeed, let $\widehat{\widehat{V}}_{\xi,n}$ be the empirical version of V_ξ, where μ_s is replaced by $\widehat{\widehat{\mu}}_{s,n}$ $(s = 1,\ldots,4)$ and by $\widehat{\mu}_{s,n}$ $(s = 5, 6)$, σ_{st} is replaced by $\widehat{\widehat{\sigma}}_{st,n}$ $(s,t \in \{1,\ldots,4\})$, σ_{sr} is replaced by $\widehat{\widehat{\sigma}}_{sr,n}$ $(s \in \{1,\ldots,4\}, r \in \{5,6\})$, and σ_{rr^*} is replaced by $\widehat{\sigma}_{rr^*,n}$ $(r,r^* \in \{5,6\})$. The consistency of $\widehat{\widehat{V}}_{\xi,n}$ is established as follows.

Theorem 4. *Under Assumptions A.1–A.10, as $n \to \infty$,*

$$\widehat{\widehat{V}}_{\xi,n} \xrightarrow{p} V_\xi. \tag{33}$$

Proof of Theorem 4. This theorem follows from Theorem 3, the fact that $\widehat{\widehat{V}}_{\xi,n}$ is obtained by replacing unknown parameters in their formulas by the corresponding consistent estimates, and the application of the continuous mapping theorem and Slutsky's theorem. □

Theorem 3 also provides a foundation to derive new central limit theorems for our target estimator, $\widehat{\widehat{\xi}}_n$. It is worth noting that this task is not a trivial adaptation of previous work (KSW2015, SZ2018, and KSW2021) because of the complicated functional form of this estimator. In particular, it involves the nonlinear operator

$\log(\cdot)$ on DEA-based components, whereas the target estimators in earlier works (e.g., SZ2018) are linear with respect to the DEA-based components weighted by different sets of weights. To circumvent this difficulty, we employ the uniform delta method (theorem 3.8 of Van der Vaart 2000), which we mention here as Lemma EC.1 in E-companion EC.2.[12] Specifically, with the help of Lemma EC.1, we can study the asymptotic behavior of sequences $\sqrt{n}(\log\widehat{\widehat{\mu}}_{s,n} - \log\widetilde{\mu}_{s,n})$ $(s = 1,2,3,4)$, where the centering vectors $\widetilde{\mu}_{s,n}$ are dependent on the sample size n as is the case in our complex statistic, and thereby we sufficiently reduce its stochastic remainder (see Remark 1 in Section 7 for further discussion). Leveraging this, we derive new central limit theorems (Theorems 5 and 6), which can then be used for a well-grounded statistical inference on aggregate MPIs.

Theorem 5. *Under Assumptions A.1–A.10, as $n \to \infty$,*

$$\sqrt{n}\left(\widehat{\widehat{\xi}}_n - \xi - C_\xi n^{-\kappa} + O(v_{n,\kappa})\right) \xrightarrow{d} \mathcal{N}(0, V_\xi), \tag{34}$$

where $C_\xi \in \mathbb{R}$ is a constant, $O(v_{n,\kappa}) = o(n^{-\kappa})$.

Because this and the next two theorems are the most important results of this paper and, to the best of our knowledge, their proofs are the most novel in the literature, we provide these proofs in the main text rather than in an appendix.

Proof of Theorem 5. For each $s = 1,2,3,4$, from Theorem 3(i), we have $\lim_{n\to\infty}\widetilde{\mu}_{s,n} = \mu_s > 0$. Hence, we can apply Lemma EC.1(i) to the result of Theorem 3(iii) and obtain

$$\begin{aligned}
\log\widehat{\widehat{\mu}}_{s,n} &= \log\widetilde{\mu}_{s,n} + \frac{1}{\mu_s}(\widehat{\widehat{\mu}}_{s,n} - \widetilde{\mu}_{s,n}) + o_p(n^{-1/2})\\
&= \log(\mu_s + C_s n^{-\kappa} + O(v_{n,\kappa})) + \frac{\widehat{\widehat{\mu}}_{s,n} - \widetilde{\mu}_{s,n}}{\mu_s} + o_p(n^{-1/2})\\
&= \log\mu_s + \frac{1}{\mu_s}[C_s n^{-\kappa} + O(v_{n,\kappa})] + O([C_s n^{-\kappa} + O(v_{n,\kappa})]^2)\\
&\quad + \frac{\widehat{\mu}_{s,n} - \mu_s + o_p(n^{-1/2})}{\mu_s} + o_p(n^{-1/2})\\
&= \log\mu_s + \frac{C_s}{\mu_s}n^{-\kappa} + O(v_{n,\kappa}) + \frac{\widehat{\mu}_{s,n} - \mu_s}{\mu_s} + o_p(n^{-1/2}).
\end{aligned} \tag{35}$$

In this expression, the third equality follows from the Taylor expansion and Theorem 3(ii), and the last equality follows from $O([C_s n^{-\kappa} + O(v_{n,\kappa})]^2) = O(n^{-2\kappa}) = o(v_{n,\kappa})$. Hereafter, we use this reasoning in subsequent proofs immediately to save space.

On the other hand, we can also apply Lemma EC.1(i) to $\sqrt{n}(\widehat{\mu}_{s,n} - \mu_s) \xrightarrow{d} \mathcal{N}(0, \sigma_{ss})$ and obtain

$$\log\widehat{\mu}_{s,n} = \log\mu_s + \frac{\widehat{\mu}_{s,n} - \mu_s}{\mu_s} + o_p(n^{-1/2}). \tag{36}$$

Note that (36) is also valid for $s = 5, 6$ because $\mu_5, \mu_6 > 0$. From (35) and (36), we have

$$\widehat{\widehat{\xi}}_n = \xi + \frac{1}{2}\left(\frac{C_3}{\mu_3} + \frac{C_4}{\mu_4} - \frac{C_1}{\mu_1} - \frac{C_2}{\mu_2}\right)n^{-\kappa} + O(v_{n,\kappa}) + o_p(n^{-1/2}) + \widehat{B}_{\xi,n}, \tag{37}$$

$$\widehat{\xi}_n = \xi + o_p(n^{-1/2}) + \widehat{B}_{\xi,n}, \tag{38}$$

where

$$\widehat{B}_{\xi,n} = \frac{\widehat{\mu}_{3,n} - \mu_3}{2\mu_3} + \frac{\widehat{\mu}_{4,n} - \mu_4}{2\mu_4} - \frac{\widehat{\mu}_{1,n} - \mu_1}{2\mu_1} - \frac{\widehat{\mu}_{2,n} - \mu_2}{2\mu_2} + \frac{\widehat{\mu}_{5,n} - \mu_5}{\mu_5} - \frac{\widehat{\mu}_{6,n} - \mu_6}{\mu_6}. \tag{39}$$

It is worth clarifying here that using the standard delta method produces a looser evaluation that cannot be used to prove Theorem 5 because a component $O_p(n^{-1/2})$ remains at the end. On the other hand, because of the uniform delta method (theorem 3.8 of Van der Vaart (2000), which we mention here as Lemma EC.1), we can extract the $O_p(n^{-1/2})$ component out (i.e., $\widehat{B}_{\xi,n}$ in (37) and (38)) so that this $O_p(n^{-1/2})$ component then cancels out when taking the difference between these two, and hence, we reduce the stochastic remainder to $o_p(n^{-1/2})$, which then helps to prove Theorem 5. Indeed, subtracting (38) from (37) yields

$$\widehat{\widehat{\xi}}_n - \widehat{\xi}_n = C_\xi n^{-\kappa} + O(v_{n,\kappa}) + o_p(n^{-1/2}), \tag{40}$$

where $C_\xi = \frac{1}{2}\left(\frac{C_3}{\mu_3} + \frac{C_4}{\mu_4} - \frac{C_1}{\mu_1} - \frac{C_2}{\mu_2}\right) \in \mathbb{R}$ is a constant.

Now, combining (B.6) from Appendix B with (40), we have the desired result. □

Theorem 5 has implications for the limiting distribution of the DEA-based estimator of ξ. In particular, $\widehat{\widehat{\xi}}_n$ is a consistent estimator of ξ with the leading bias term being $C_\xi n^{-\kappa}$. Moreover, the asymptotic behavior of this bias when multiplied with the norming rate $\sqrt{n}$ is revealed as follows.

- If $\kappa > 1/2$ $(p + q = 2)$, the bias term in (34) vanishes asymptotically and can be ignored.
- If $\kappa = 1/2$ $(p + q = 3)$, the bias term converges to an unknown constant, implying that Theorem 5 cannot be used immediately to make inferences about ξ.
- If $\kappa < 1/2$ $(p + q = 4, 5, 6, \ldots)$, the bias term explodes to infinity as n increases, and again, Theorem 5 cannot be used directly to make inferences about the parameter of interest.

As a consequence, there emerges a need to correct for the bias term in Theorem 5 in order to make inferences when $\kappa \le 1/2$. In the spirit of KSW2015, SZ2018, and KSW2021, we find another norming rate different from $\sqrt{n}$ for the case $\kappa \le 1/2$. Specifically, let $\widetilde{n}$ denote the appropriately adjusted sample size that can be determined as[13]

$$\widetilde{n} := \widetilde{n}_\kappa = \min\{\lfloor n^{2\kappa} \rfloor, n\} \le n. \quad (41)$$

Now, consider the estimator $\widehat{\widehat{\xi}}_{\tilde{n}}$, which is the subsample version of $\widehat{\widehat{\xi}}_n$ in the sense that the averages are taken over a random subsample $\mathcal{X}^*_{\tilde{n}}$ of size $\widetilde{n}$, where $\mathcal{X}^*_{\tilde{n}} \subset \mathcal{X}_n$. Formally,

$$\begin{aligned}\widehat{\widehat{\xi}}_{\tilde{n}} = \frac{1}{2}\Bigg[&\log\left(\frac{\sum_{\{i:Z_i^1\in\mathcal{X}^*_{\tilde{n}}\}}\lambda_C(Z_i^1|\mathcal{X}_n^1)w^1Y_i^1}{\widetilde{n}}\right)\\ &+\log\left(\frac{\sum_{\{i:Z_i^1\in\mathcal{X}^*_{\tilde{n}}\}}\lambda_C(Z_i^1|\mathcal{X}_n^2)w^1Y_i^1}{\widetilde{n}}\right)\\ &-\log\left(\frac{\sum_{\{i:Z_i^2\in\mathcal{X}^*_{\tilde{n}}\}}\lambda_C(Z_i^2|\mathcal{X}_n^1)w^2Y_i^2}{\widetilde{n}}\right)\\ &-\log\left(\frac{\sum_{\{i:Z_i^2\in\mathcal{X}^*_{\tilde{n}}\}}\lambda_C(Z_i^2|\mathcal{X}_n^2)w^2Y_i^2}{\widetilde{n}}\right)\Bigg]\\ &+\log\left(\frac{\sum_{\{i:Z_i^2\in\mathcal{X}^*_{\tilde{n}}\}}w^2Y_i^2}{\widetilde{n}}\right)-\log\left(\frac{\sum_{\{i:Z_i^1\in\mathcal{X}^*_{\tilde{n}}\}}w^1Y_i^1}{\widetilde{n}}\right).\end{aligned} \quad (42)$$

Similar to KSW2015, SZ2018, and KSW2021, note that, although the average in (42) is taken over the subsample $\mathcal{X}^*_{\tilde{n}}$, the DEA efficiency scores are still estimated using all of the available observations in the original sample $\mathcal{X}_n$. The asymptotics of this new estimator are revealed as follows.

Theorem 6. *Under Assumptions A.1–A.10, when $\kappa \le 1/2$ as $n \to \infty$,*

$$n^\kappa\left(\widehat{\widehat{\xi}}_{\tilde{n}} - \xi - C_\xi n^{-\kappa} + O(v_{n,\kappa})\right) \xrightarrow{d} \mathcal{N}(0, V_\xi), \quad (43)$$

where C_ξ is the same constant as in Theorem 5 and $O(v_{n,\kappa}) = o(n^{-\kappa})$.

Proof of Theorem 6. For convenience, we can assume that the observations in $\mathcal{X}_n$ are randomly sorted and $\mathcal{X}^*_{\tilde{n}}$ consists of the first $\widetilde{n}$ elements of the sorted sample. Before going into details of the proof, we establish some asymptotic properties similar to Theorem 3 as follows.

For $s = 1, \ldots, 4$, let $\chi_{s,\tilde{n}} = \widehat{\widehat{\mu}}_{s,\tilde{n}} - \widehat{\mu}_{s,\tilde{n}}$, where $\widehat{\widehat{\mu}}_{s,\tilde{n}} = \widetilde{n}^{-1}\sum_{i=1}^{\tilde{n}} \widehat{U}_{s,i}$ and $\widehat{\mu}_{s,\tilde{n}} = \widetilde{n}^{-1}\sum_{i=1}^{\tilde{n}} U_{s,i}$. Remember that DEA estimation in $\widehat{U}_{s,i}$ here is the same as before, that is, using all observations in the original sample $\mathcal{X}_n$. In addition, let $\widetilde{\mu}_{s,\tilde{n}} = E(\widehat{\widehat{\mu}}_{s,\tilde{n}})$. Then, by Theorem 1, we have

$$\begin{aligned}E(\chi_{s,\tilde{n}}) = \widetilde{\mu}_{s,\tilde{n}} - \mu_s &= \widetilde{n}^{-1}\sum_{i=1}^{\tilde{n}} E(\widehat{U}_{s,i} - U_{s,i})\\ &= C_s n^{-\kappa} + O(v_{n,\kappa}),\end{aligned} \quad (44)$$

and

$$\begin{aligned}Var(\chi_{s,\tilde{n}}) &= \widetilde{n}^{-2}\sum_{i=1}^{\tilde{n}} Var(\widehat{U}_{s,i} - U_{s,i})\\ &= \widetilde{n}^{-2}\sum_{i=1}^{\tilde{n}}(E([\widehat{U}_{s,i} - U_{s,i}]^2) - (E(\widehat{U}_{s,i} - U_{s,i}))^2)\\ &= \widetilde{n}^{-1}(o(n^{-\kappa}) - (C_s n^{-\kappa} + O(v_{n,\kappa}))^2)\\ &= \widetilde{n}^{-1}o(n^{-\kappa}),\end{aligned} \quad (45)$$

where C_s is the same constant as in Theorem 3 ($s = 1, \ldots, 4$).

Consequently, by Markov's inequality, for any $\epsilon > 0$, we have

$$\begin{aligned}Pr\left(\sqrt{\widetilde{n}}|\chi_{s,\tilde{n}} - E(\chi_{s,\tilde{n}})| > \epsilon\right) &\le \frac{E(\widetilde{n}(\chi_{s,\tilde{n}} - E(\chi_{s,\tilde{n}}))^2)}{\epsilon^2}\\ &= \frac{\widetilde{n}Var(\chi_{s,\tilde{n}})}{\epsilon^2} = \frac{o(n^{-\kappa})}{\epsilon^2},\end{aligned}$$

which implies that $\chi_{s,\tilde{n}} - E(\chi_{s,\tilde{n}}) = o_p(\widetilde{n}^{-1/2})$ or, equivalently,

$$\widehat{\widehat{\mu}}_{s,\tilde{n}} - \widetilde{\mu}_{s,\tilde{n}} = \widehat{\mu}_{s,\tilde{n}} - \mu_s + o_p(\widetilde{n}^{-1/2}), \quad (46)$$

and as a consequence,

$$\sqrt{\widetilde{n}}\left(\widehat{\widehat{\mu}}_{s,\tilde{n}} - \widetilde{\mu}_{s,\tilde{n}}\right) \xrightarrow{d} \mathcal{N}(0, \sigma_{ss}). \quad (47)$$

By virtue of (44), the equality can also be presented as

$$\begin{aligned}\widehat{\widehat{\mu}}_{s,\tilde{n}} - \widehat{\mu}_{s,\tilde{n}} &= \widetilde{\mu}_{s,\tilde{n}} - \mu_s + o_p(\widetilde{n}^{-1/2})\\ &= C_s n^{-\kappa} + O(v_{n,\kappa}) + o_p(\widetilde{n}^{-1/2}),\\ &\qquad s = 1, \ldots, 4.\end{aligned} \quad (48)$$

Because $\kappa \le 1/2$, $\widetilde{n} = \lfloor n^{2\kappa} \rfloor \le n$ and $\lim_{n\to\infty}\frac{n^\kappa}{\sqrt{\tilde{n}}} = \lim_{n\to\infty}\frac{n^\kappa}{\sqrt{\lfloor n^{2\kappa}\rfloor}} = 1$, it is sufficient to prove that

$$\sqrt{\widetilde{n}}\left(\widehat{\widehat{\xi}}_{\tilde{n}} - \xi - C_\xi n^{-\kappa} + O(v_{n,\kappa})\right) \xrightarrow{d} \mathcal{N}(0, V_\xi), \text{ as } n \to \infty. \quad (49)$$

From (44), we have $\lim_{\tilde{n}\to\infty}\widetilde{\mu}_{s,\tilde{n}} = \mu_s > 0$ for $s = 1,2,3,4$. Thus, similar to the proof of Theorem 5, we can apply Lemma EC.1(i) to (47) and then use a Taylor expansion to get the following result:

$$\begin{aligned}\log\widehat{\widehat{\mu}}_{s,\tilde{n}} &= \log\widetilde{\mu}_{s,\tilde{n}} + \frac{1}{\mu_s}(\widehat{\widehat{\mu}}_{s,\tilde{n}} - \widetilde{\mu}_{s,\tilde{n}}) + o_p(\widetilde{n}^{-1/2})\\ &= \log(\mu_s + C_s n^{-\kappa} + O(v_{n,\kappa})) + \frac{\widehat{\mu}_{s,\tilde{n}} - \mu_s + o_p(\widetilde{n}^{-1/2})}{\mu_s}\\ &\quad + o_p(\widetilde{n}^{-1/2})\\ &= \log\mu_s + \frac{1}{\mu_s}(C_s n^{-\kappa} + O(v_{n,\kappa})) + O([C_s n^{-\kappa} + O(v_{n,\kappa})]^2)\\ &\quad + \frac{\widehat{\mu}_{s,\tilde{n}} - \mu_s}{\mu_s} + o_p(\widetilde{n}^{-1/2})\\ &= \log\mu_s + \frac{C_s}{\mu_s}n^{-\kappa} + O(v_{n,\kappa}) + \frac{\widehat{\mu}_{s,\tilde{n}} - \mu_s}{\mu_s} + o_p(\widetilde{n}^{-1/2}).\end{aligned} \quad (50)$$

On the other hand, it can be deduced from (36) that

$$\log \widehat{\mu}_{s,\tilde{n}} = \log \mu_s + \frac{\widehat{\mu}_{s,\tilde{n}} - \mu_s}{\mu_s} + o_p(\tilde{n}^{-1/2}). \tag{51}$$

Subtracting (51) from (50) yields

$$\log \widehat{\widehat{\mu}}_{s,\tilde{n}} - \log \widehat{\mu}_{s,\tilde{n}} = \frac{C_s}{\mu_s} n^{-\kappa} + O(v_{n,\kappa}) + o_p(\tilde{n}^{-1/2}) \quad (s = 1,2,3,4). \tag{52}$$

Therefore, we obtain

$$\widehat{\widehat{\xi}}_{\tilde{n}} - \widehat{\xi}_{\tilde{n}} = C_\xi n^{-\kappa} + O(v_{n,\kappa}) + o_p(\tilde{n}^{-1/2}), \tag{53}$$

where $C_\xi = \frac{1}{2}\left(\frac{C_3}{\mu_3} + \frac{C_4}{\mu_4} - \frac{C_1}{\mu_1} - \frac{C_2}{\mu_2}\right)$ is the same constant as in Theorem 5, and $\widehat{\xi}_{\tilde{n}}$ is the analogue of $\widehat{\xi}_n$ but computed using the subsample $\mathcal{X}^*_{\tilde{n}}$ (note that there is no DEA estimation in $\widehat{\xi}_{\tilde{n}}$). Clearly, this result, when combined with $\sqrt{\tilde{n}}(\widehat{\xi}_{\tilde{n}} - \xi) \xrightarrow{d} \mathcal{N}(0, V_\xi)$ as $\tilde{n} \to \infty$, leads to the desired result. □

Intuitively, and in a nutshell, Theorem 6 says that, when $\kappa < 1/2$, the bias term converges to a constant instead of exploding to infinity as was demonstrated in Theorem 5.[14] Theorems 5 and 6 provide the foundation for a well-grounded statistical inference on aggregate MPIs. To implement them, we need suitable estimators of the bias, and we discuss these in the next section.

5. Generalized Jackknife Estimators for the Bias

In the spirit of KSW2015, SZ2018, and KSW2021, we construct a jackknife-type bias estimator on the basis of splitting the original sample into two subsamples. The details are presented subsequently, and one may notice that the adaptation of the previous works in the related literature is nontrivial.

For each $l = 1, \ldots, L$ where $L \ll \binom{n}{\lfloor n/2 \rfloor}$, split $\mathcal{X}_n$ randomly into two subsamples $\mathcal{X}_{l,m_1}$ and $\mathcal{X}_{l,m_2}$ of sizes $m_1 = \lfloor n/2 \rfloor$ and $m_2 = n - m_1$, respectively. More precisely, $\mathcal{X}_n = \mathcal{X}_{l,m_1} \cup \mathcal{X}_{l,m_2}$, where, for each $j = 1, 2$, $\mathcal{X}_{l,m_j} = \mathcal{X}^1_{l,m_j} \cup \mathcal{X}^2_{l,m_j}$ consists of sets of the same m_j DMUs observed in periods 1 and 2, that is, $\mathcal{X}^1_{l,m_j}$ and $\mathcal{X}^2_{l,m_j}$, respectively. Then, for each $j = 1, 2$, set

$$\widehat{\widehat{\mu}}_{l,1,m_j} = m_j^{-1} \sum_{\{i: Z_i^2 \in \mathcal{X}^2_{l,m_j}\}} \lambda_C(Z_i^2 \mid \mathcal{X}^1_{l,m_j}) w^2 Y_i^2, \tag{54}$$

and similarly, define $\widehat{\widehat{\mu}}_{l,s,m_j}$ $(s = 2,3,4)$ as the analogues of $\widehat{\widehat{\mu}}_{s,n}$ in the same way as $\widehat{\widehat{\mu}}_{l,1,m_j}$. Here, it is important to highlight that, unlike $\widehat{\widehat{\xi}}_{\tilde{n}}$, the efficiency scores in $\widehat{\widehat{\mu}}_{l,s,m_j}$ are estimated by DEA using only observations in the corresponding subsample $\mathcal{X}_{l,m_j}$. For $j = 1, 2$, we also define

$$\widehat{\widehat{\xi}}_{l,m_j} = \frac{1}{2}\left(\log \widehat{\widehat{\mu}}_{l,3,m_j} + \log \widehat{\widehat{\mu}}_{l,4,m_j} - \log \widehat{\widehat{\mu}}_{l,1,m_j} - \log \widehat{\widehat{\mu}}_{l,2,m_j}\right) + \log \widehat{\mu}_{5,n} - \log \widehat{\mu}_{6,n}. \tag{55}$$

In essence, $\widehat{\widehat{\xi}}_{l,m_j}$ is an analogue of $\widehat{\widehat{\xi}}_n$ in the sense that components containing efficiency scores are evaluated over the subsample $\mathcal{X}_{l,m_j}$, whereas the other components (i.e., $\widehat{\mu}_{5,n}$ and $\widehat{\mu}_{6,n}$) are evaluated over the original full sample $\mathcal{X}_n$. Now, let

$$\widehat{\widehat{\xi}}^*_{l,n} = \frac{1}{2}\left(\widehat{\widehat{\xi}}_{l,m_1} + \widehat{\widehat{\xi}}_{l,m_2}\right).$$

The generalized jackknife estimator of the bias associated with $\widehat{\widehat{\xi}}_n$ is given by

$$\widehat{A}_{\xi,n,\kappa,L} = L^{-1} \sum_{l=1}^{L} (2^\kappa - 1)^{-1} \left(\widehat{\widehat{\xi}}^*_{l,n} - \widehat{\widehat{\xi}}_n\right).$$

The following theorem reveals an important asymptotic property of this bias estimator.

Theorem 7. *Under Assumptions A.1–A.10, as $n \to \infty$,*

$$\widehat{A}_{\xi,n,\kappa,L} = C_\xi n^{-\kappa} + O(v_{n,\kappa}) + o_p(n^{-1/2}), \tag{56}$$

where C_ξ is the same constant as in Theorems 5 and 6, and $O(v_{n,\kappa}) = o(n^{-\kappa})$.

Proof of Theorem 7. If n is odd, $\frac{m_1}{m_2} = \frac{\lfloor n/2 \rfloor}{\lfloor n/2 \rfloor + 1} \to 1$ as $n \to \infty$. Thus, for simplicity, we can assume without affecting the asymptotical result of the theorem that n is even and $m_1 = m_2 = n/2$.

For each $l = 1, \ldots, L$, $s = 1,2,3,4$ and $j = 1, 2$, it follows by the same arguments in the proof of Theorem 5 that

$$\begin{aligned}\widehat{\widehat{\xi}}_{l,m_j} &= \xi + C_\xi (n/2)^{-\kappa} + O(v_{n/2,\kappa}) + \widehat{B}_{\xi,l,m_j} + o_p(n^{-1/2}) \\ &= \xi + 2^\kappa C_\xi n^{-\kappa} + O(v_{n,\kappa}) + \widehat{B}_{\xi,l,m_j} + o_p(n^{-1/2}),\end{aligned} \tag{57}$$

where $\widehat{B}_{\xi,l,m_j}$ is the analogue of $\widehat{B}_{\xi,n}$ in the sense that components involving efficiency scores are evaluated over the subsample $\mathcal{X}_{l,m_j}$, whereas the other components (i.e., $\widehat{\mu}_{5,n}$ and $\widehat{\mu}_{6,n}$) are evaluated over the full sample $\mathcal{X}_n$. Formally,

$$\widehat{B}_{\xi,l,m_j} = \frac{\widehat{\mu}_{l,3,m_j} - \mu_3}{2\mu_3} + \frac{\widehat{\mu}_{l,4,m_j} - \mu_4}{2\mu_4} - \frac{\widehat{\mu}_{l,1,m_j} - \mu_1}{2\mu_1} - \frac{\widehat{\mu}_{l,2,m_j} - \mu_2}{2\mu_2} + \frac{\widehat{\mu}_{5,n} - \mu_5}{\mu_5} - \frac{\widehat{\mu}_{6,n} - \mu_6}{\mu_6},$$

where $\widehat{\mu}_{l,s,m_j}$ is the analogue of $\widehat{\mu}_{s,n}$ but evaluated over the subsample $\mathcal{X}_{l,m_j}$. Note that, in (57), we use the fact that $o_p((n/2)^{-1/2}) = o_p(n^{-1/2})$ and $O(v_{n/2,\kappa}) = O(v_{n,\kappa})$, which is because of

$$\lim_{n\to\infty}\frac{v_{n/2,\kappa}}{v_{n,\kappa}}=\lim_{n\to\infty}\frac{\left(\frac{n}{2}\right)^{-\frac{3}{p+q+1}}\left(\log\frac{n}{2}\right)^{\frac{3}{p+q+1}}}{n^{-\frac{3}{p+q+1}}(\log n)^{\frac{3}{p+q+1}}}$$
$$=\lim_{n\to\infty}2^{\frac{3}{p+q+1}}\left(1-\frac{\log 2}{\log n}\right)^{\frac{3}{p+q+1}}=2^{\frac{3}{p+q+1}}. \quad (58)$$

Taking the average of (57) over $j = 1, 2$, we can come up with

$$\widehat{\widehat{\xi}}^{\,*}_{l,n}=\xi+2^{\kappa}C_{\xi}n^{-\kappa}+O(v_{n,\kappa})+\frac{1}{2}(\widehat{B}_{\xi,l,m_1}+\widehat{B}_{\xi,l,m_2})+o_p(n^{-1/2}). \quad (59)$$

Subtracting (37) from this equality yields

$$\widehat{\widehat{\xi}}^{\,*}_{l,n}-\widehat{\widehat{\xi}}_n=(2^{\kappa}-1)C_{\xi}n^{-\kappa}+O(v_{n,\kappa})+\left[\frac{1}{2}(\widehat{B}_{\xi,l,m_1}+\widehat{B}_{\xi,l,m_2})-\widehat{B}_{\xi,n}\right]+o_p(n^{-1/2})$$
$$=(2^{\kappa}-1)C_{\xi}n^{-\kappa}+O(v_{n,\kappa})+o_p(n^{-1/2}). \quad (60)$$

In this expression, the second equality follows from $\frac{1}{2}(\widehat{B}_{\xi,l,m_1}+\widehat{B}_{\xi,l,m_2})-\widehat{B}_{\xi,n}=0$, which is because, for $s=1,\ldots,6$,

$$\widehat{\mu}_{s,n}-\frac{1}{2}(\widehat{\mu}_{l,s,m_1}+\widehat{\mu}_{l,s,m_2})$$
$$=\frac{1}{n}\sum_{i=1}^{n}U_{s,i}-\frac{1}{2}\left(\frac{1}{n/2}\sum_{\{i:Z_i^t\in\mathcal{X}_{l,m_1}\}}U_{s,i}+\frac{1}{n/2}\sum_{\{i:Z_i^t\in\mathcal{X}_{l,m_2}\}}U_{s,i}\right)$$
$$=\frac{1}{n}\sum_{i=1}^{n}U_{s,i}-\frac{1}{n}\sum_{i=1}^{n}U_{s,i}=0. \quad (60)$$

Finally, taking the average of (60) over $l=1,\ldots,L$ leads to the desired result. □

It should be emphasized that, under Assumptions A.1–A.10, the variance of $\widehat{A}_{\xi,n,\kappa,L}$ is inversely proportional to L^2. Therefore, similar to Kneip et al. (2016), whereas Theorem 7 holds true even with $L=1$, applied researchers might want to increase L (e.g., $L=100$) to reduce the variance of this bias estimator and achieve more reliable bias corrections for a few particular samples in practice.[15]

6. Confidence Intervals

As discussed in Section 4, statistical inferences for ξ can be obtained directly from Theorem 5 when $\kappa>1/2$ (i.e., $p+q=2$) by ignoring the bias because it disappears asymptotically when multiplied with the norming rate $\sqrt{n}$. However, it might not be ideal to do so in practice because the bias might still be significant in small samples. In light of Theorem 7, we can account for this issue by estimating the leading term of the bias. The following theorem presents important results that pave the way for making inferences about the parameter of interest ξ for all cases of κ.

Theorem 8. *Under Assumptions A.1–A.10, as $n\to\infty$ for $\kappa\geq 1/2$,*

$$\sqrt{n}\left(\widehat{\widehat{\xi}}_n-\xi-\widehat{A}_{\xi,n,\kappa,L}+O(v_{n,\kappa})\right)\xrightarrow{d}\mathcal{N}(0,V_{\xi}), \quad (62)$$

and for $0<\kappa<1/2$,

$$n^{\kappa}\left(\widehat{\widehat{\xi}}_{\tilde{n}}-\xi-\widehat{A}_{\xi,n,\kappa,L}+O(v_{n,\kappa})\right)\xrightarrow{d}\mathcal{N}(0,V_{\xi}), \quad (63)$$

where $\kappa=2/(p+q+1)$ and $O(v_{n,\kappa})=o(n^{-\kappa})$.

This theorem follows from substituting the results from Theorem 7 into the corresponding ones in Theorem 5, noting that $\sqrt{n}o_p(n^{-1/2})=o_p(1)$ and $n^{\kappa}o_p(n^{-1/2})=o_p(1)$ when $\kappa<1/2$.

In light of Theorems 4 and 8 and Slutsky's theorem, feasible confidence intervals for ξ can now be derived straightforwardly with the note that $O(v_{n,\kappa})$ in Theorem 8 can be ignored because $\sqrt{n}O(v_{n,\kappa})=\sqrt{n}o(n^{-\kappa})=o(1)$ when $\kappa\geq 1/2$ and $n^{\kappa}O(v_{n,\kappa})=n^{\kappa}o(n^{-\kappa})=o(1)$ when $\kappa<1/2$. In particular, asymptotically correct $100(1-\alpha)\%$ symmetric confidence intervals for ξ are given by

$$\left[\widehat{\widehat{\xi}}_n-\widehat{A}_{\xi,n,\kappa,L}\pm\Phi^{-1}_{1-\alpha/2}\sqrt{\widehat{\widehat{V}}_{\xi,n}/n}\right], \quad (64)$$

$$\left[\widehat{\widehat{\xi}}_{\tilde{n}}-\widehat{A}_{\xi,n,\kappa,L}\pm\Phi^{-1}_{1-\alpha/2}\sqrt{\widehat{\widehat{V}}_{\xi,n}/n^{\kappa}}\right], \quad (65)$$

for $\kappa\geq 1/2$ (i.e., $p+q=2,3$) and $\kappa<1/2$ (i.e., $p+q=4,5,6,\ldots$), respectively. Here, we recall $\tilde{n}=\min\{\lfloor n^{2\kappa}\rfloor,n\}$, and $\Phi^{-1}_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of the standard normal distribution.

7. Some Important Remarks

A few important clarifications regarding the theoretical results derived earlier are in order.

Remark 1. The theoretical developments in this paper are substantial (and nontrivial) generalizations of previous works, particularly SZ2018. The major difficulty here relative to SZ2018 is that the functional form of the aggregate MPIs (i.e., weighted harmonic-type mean) is much more complex than the form for the aggregate Farrell-type efficiency measures: the former includes nonlinear operators on DEA-based estimates entering the MPI via different components, each with a different weighting scheme. As such, arguments for deriving limiting distributions and bias corrections in SZ2018 cannot be trivially adapted to the aggregate MPIs. To overcome this complexity, we express the DEA-based estimators of ξ as sums of (i) the underlying parameter of interest, (ii) the bias term, (iii) a stochastic term of an order smaller than $n^{-1/2}$, and (iv) an expression that is linear with respect to DEA-based components (e.g., see Equations (37) and (59)). Then, we deploy the uniform delta method (i.e., Lemma EC.1 or, equivalently, theorem 3.8 of Van der Vaart 2000) to prove the asymptotic behavior of sequences $\sqrt{n}(\log\widehat{\mu}_{s,n} - \log\widetilde{\mu}_{s,n})$ $(s = 1,2,3,4)$, where, unlike in standard cases, the centering vectors $\widetilde{\mu}_{s,n}$ are dependent on the sample size n. It is also worth noting here again that the standard delta method produces a looser evaluation that cannot prove Theorem 5 because the stochastic remainder is $O_p(n^{-1/2})$. On the other hand, deploying the uniform delta method helps in canceling out the $O_p(n^{-1/2})$ term, thus reducing the stochastic remainder to $o_p(n^{-1/2})$ as shown by Equation (40). Based on this, the task of deriving the limiting distributions (Theorems 5 and 6) and the jackknife bias correction can be carried out smoothly. To the best of our knowledge, this strategy of the proof is novel relative to the previous works in this area.

Remark 2. Similar to KSW2015, SZ2018, and KSW2021, although our theory suggests using two different confidence intervals corresponding to $\kappa \geq 1/2$ and $0 < \kappa < 1/2$, it is worth noting that the former is still valid for $\kappa = 2/5$ (i.e., $p + q = 4$). Indeed, when $\kappa = 2/5$, we have

$$\sqrt{n}O(v_{n,\kappa}) = \sqrt{n}O\left(\left(\frac{\log n}{n}\right)^{3/5}\right) = O\left(\frac{(\log n)^{3/5}}{n^{1/10}}\right) = o(1), \tag{66}$$

$$n^{\kappa}O(v_{n,\kappa}) = n^{2/5}O\left(\left(\frac{\log n}{n}\right)^{3/5}\right) = O\left(\frac{(\log n)^{3/5}}{n^{1/5}}\right) = o(1), \tag{67}$$

and hence, the bias term $O(v_{n,\kappa})$ disappears asymptotically with both norming rates $\sqrt{n}$ and n^{κ} in both confidence intervals. However, one may notice from these equalities that $n^{\kappa}O(v_{n,\kappa})$ converges to zero faster than $\sqrt{n}O(v_{n,\kappa})$, which fortifies the use of the confidence interval with norming rate n^{κ} for $\kappa = 2/5 < 1/2$ as before. We also check and confirm this remark in our Monte Carlo experiments (E-companion EC.3.2).

Remark 3. As noted by KSW2015 and KSW2021, when $0 < \kappa < 1/2$, one can obtain more informative confidence intervals by employing a recentering technique. To do so, one needs to replace the point estimator using only $\widetilde{n}$ DMUs (i.e., $\widehat{\widehat{\xi}}_{\widetilde{n}}$) by its analogue evaluated over the full original sample (i.e., $\widehat{\widehat{\xi}}_{n}$). The recentered version of confidence interval (65) is as follows:

$$\left[\widehat{\widehat{\xi}}_{n} - \widehat{A}_{\xi,n,\kappa,L} \pm \Phi^{-1}_{1-\alpha/2}\sqrt{\widehat{\widehat{V}}_{\xi,n}}/n^{\kappa}\right]. \tag{68}$$

Similar to KSW2015 and KSW2021, this recentering technique helps average the center over all possible draws (without replacement) of subsamples of size $\widetilde{n}$ and, hence, eliminates the randomness as well as any deviation because of calculation on only a random subset of $\widetilde{n}$ DMUs. In fact, one can see, for example, that $\widehat{\widehat{\xi}}_{n}$ is a better estimator of ξ than $\widehat{\widehat{\xi}}_{\widetilde{n}}$ in terms of mean-square error. Moreover, the coverage of the recentered confidence interval converges to one as $n \to \infty$, exhibiting greater coverage and having the same width as the respective origin whose coverage converges to $1 - \alpha$ as $n \to \infty$.[16] This is confirmed by the Monte Carlo evidence in E-companion EC.3.2.

Remark 4. It is possible that the bias cancels out in certain situations, for example, $C_{\xi} = \frac{1}{2}\left(\frac{C_3}{\mu_3} + \frac{C_4}{\mu_4} - \frac{C_1}{\mu_1} - \frac{C_2}{\mu_2}\right) = 0$. As a result, statistical inferences using a naive application of the standard central limit theorem (i.e., using the confidence interval (B.11)) with unobserved elements being replaced by their respective DEA estimates) might still be valid under these circumstances.

KSW2021 mentions the possibility of having no bias in the simple Malmquist indices and shows that, under very peculiar (and maybe unrealistic) assumptions on the data-generating process, this bias may be equal to zero (see KSW2021, theorems 3.5 and 3.6 and remark 3.1). In particular, some necessary (but not sufficient) conditions for the bias to cancel out are that (i) two samples are generated from identical data-generating processes or (ii) the joint density of input–output bundles in two time periods is symmetric.

Remark 5. We elaborate on Remark 4 and note that, even if the bias does not cancel out, in some peculiar cases, it might still be so tiny that its theoretical explosion to infinity implied from Theorem 5 is not clearly observed in moderate sample sizes and low dimensions (i.e., the number of inputs and outputs). As such, the naive application of the standard central limit theorem might also perform fairly well in this context. For example, in Theorem 5, if $C_{\xi} = 0.01$ and $\kappa = 1/3$ (i.e., $p + q = 5$), then the leading bias term $C_{\xi}n^{-\kappa}$ multiplied with the norming rate $\sqrt{n}$ is equal to 0.031, 0.046, 0.068 for $n = 1{,}000$, $10{,}000$, $100{,}000$, respectively.

Remark 6. In addition to Remark 5, it can be seen that the coefficient of the leading bias term in this paper is obtained from subtracting symmetric terms. More precisely, in C_{ξ}, the constant C_s $(s = 1,2,3,4)$ is normalized by the respective population means μ_s (i.e., $\frac{C_s}{\mu_s}$). Thus, in practice, the chance that the bias coefficient C_{ξ} is close to zero (or even cancels out), making the whole bias negligible in moderate sample sizes and low dimensions, might be relatively higher than that for KSW2015, SZ2018, and KSW2021.

Because the magnitude of the bias is unknown in reality, the naive application of the standard central limit theorems in dimensions greater than two (i.e., $\kappa \leq 1/2$) is problematic as pointed out in Theorem 5. As such, the use of the theoretically justified theorems appears to be a safer choice.

Remark 7. The evidence from extensive Monte Carlo experiments that we perform for various scenarios supports the developed asymptotic theory and the aforementioned Remarks 1–6 in particular, and is summarized in Appendix EC.3.

8. Empirical Illustration

One of the most popular applications of DEA in general and the MPI in particular is analyzing the performance of countries relative to what is sometimes referred to as the estimated best practice world technology frontier. In fact, one of the first applications of MPI is also in this area from the seminal work of Färe et al. (1994), which, in turn, inspired many other studies in the literature. Among these studies are Kumar and Russell (2002) with a sample of 57 countries over the period 1965–1990 extracted from the Penn World Tables (version 5.6) and Badunenko et al. (2008) with a sample of 84 countries over the period 1990–2000, and here, we consider these same 84 countries over the period 1990–2019 from Penn World Tables version 10.0 (PWT10).

Briefly, our aggregate measure of output is the output-side real GDP at chained purchasing power parity (PPP) rates (in millions 2017US$) (i.e., *rgdpo* in the PWT10), our aggregate measures of inputs (capital and labor) are capital stock at current PPPs (in millions 2017US$) and the number of persons engaged (in millions) (i.e., *cn* and *emp* in the PWT10, respectively).

The estimation results for PWT10 are presented in Table 1, in which we consider three groups over which to aggregate: (i) the entire sample, (ii) developed countries, and (iii) developing countries.[17] For each group, we consider two approaches: (a) the weighted MPI framework developed in this paper, that is, in which the weight of each country is accounted for in the averaging of the estimated indexes, and (b) when just the sample mean (i.e., equal weighting) of the estimated indexes is used. The first three rows of numbers in Table 1 present the former approach, whereas the last three rows of numbers present the latter approach. Also note that the first two columns of numbers in Table 1 present the estimates before and after the bias correction, respectively, whereas the last two columns present the lower and upper bounds for the 95% confidence intervals (constructed around the bias-corrected estimates). The results largely speak for themselves, yet a few brief remarks are worth highlighting.

First, we see fairly substantial differences between the weighted and nonweighted approaches. Indeed, when we compare the bias-corrected estimates for the entire sample, the weighted approach yields an aggregate index of about 0.93 (i.e., implying 7% deterioration), whereas the nonweighted approach gives about 1.06 (i.e., 6% growth), thus suggesting about 13% higher growth estimate. Interestingly, note that a much smaller difference is observed when comparing the estimates without the bias correction developed in this paper: 0.99 versus 1.04.

Next, an interesting picture is also seen when focusing on the subset of developed countries ($n = 27$). Indeed, note that, using the bias-corrected estimates, the weighted approach gives about 1.16, whereas the nonweighted approach gives 1.05. Meanwhile, the original estimates (before the bias correction) are about 1.16 versus 1.13. That is, we see another illustration of the importance of deploying the bias correction for the MPIs: the bias-corrected estimate shows lower aggregate growth of the group of developed countries when using the nonweighted averaging of MPI.

An interesting picture also arises when we focus on the subset of developing countries ($n = 57$). Indeed, here, when we compare the bias-corrected estimates, the nonweighted approach gives 1.18 (i.e., a growth of 18%), whereas the weighted approach gives about 1.05 (i.e., a growth of only 5%). Meanwhile, note that the difference is much smaller in absolute terms when no bias correction is deployed (1.11 versus 1.13), hence illustrating the importance of the bias correction. Moreover, we see that, in the nonweighted approach, the estimated confidence interval does not cover unity, whereas the estimated confidence interval for the weighted approach covers unity.

Table 1. Estimation Results and 95% Confidence Intervals for a Sample from PWT10

	$\exp\left(\widehat{\overline{\xi}}_{\tilde{n}}\right)$	$\exp\left(\widehat{\overline{\xi}}_{\tilde{n}} - \widehat{A}_{\xi,n,\kappa,L}\right)$	Lower bound	Upper bound
Entire sample	0.9850	0.9347	0.7961	1.0976
Developed	1.1570	1.1563	1.0266	1.3023
Developing	1.1297	1.0523	0.9542	1.1605
	$\exp(\widehat{E}_{\tilde{n}}(\log(M_O)))$	$\exp(\widehat{E}_{\tilde{n}}(\log(M_O)) - \widehat{B}_{n,\kappa,L})$	Lower bound	Upper bound
Entire sample	1.0388	1.0640	0.9658	1.1721
Developed	1.1284	1.0538	0.9769	1.1367
Developing	1.1114	1.1782	1.0202	1.3607

Note. Number of random splits $L = 1{,}000$; $\kappa = 2/(p + q + 1) = 0.5$; $n = 27, 57, 84$ for developed countries, developing countries, and entire sample, respectively.

All in all, this empirical example illustrates the importance of considering the aggregation that accounts for an economic weight associated with an index in addition to the simple averaging (because the two approaches can imply very different conclusions and support different policy implications) as well as the importance of the use of the new central limit theorems (with bias correction) for estimating confidence intervals for the weighted MPIs.

9. Concluding Remarks

Whereas it is easy to see that accounting for economic weights of individuals in aggregations of indices makes sense and it may yield very different conclusions compared with the equal-weight aggregations, its practical implementation still needs some statistical theory (for constructing confidence intervals and performing hypothesis tests), which have never been done before. In this paper, we fill this gap by developing a comprehensive set of asymptotic properties for the meaningful aggregation of MPIs, a weighted harmonic-type mean of individual efficiency scores. This provides operational researchers with the tools for constructing theoretically justified confidence intervals and statistical tests for the aggregate productivity indices. Our Monte Carlo evidence confirms that the newly developed statistical inferences perform well in finite samples similar to that in KSW2015, SZ2018, and KSW2021.

Our empirical application to real data also vividly illustrate the importance of considering the weighted (in addition to nonweighted) aggregation of productivity indexes, especially with the theoretically justified bias correction and confidence intervals developed in this paper. For example, the weighted approach gives substantially different estimates of aggregate growth (and corresponding conclusions) compared with the nonweighted approach before and after the bias correction as well as fairly different estimates of confidence intervals.

It is also worth noting that we contribute a new approach in this literature for deriving asymptotic properties for complex indices that are nonlinear with respect to their DEA-based components. This approach paves the way for deriving a similar theory for other sophisticated indices, such as the Hicks–Moorsteen productivity index, although this would be a study in itself because of the even greater complexity of this index. This paper also provides important statistical grounds for further theoretical developments. For instance, one may develop a test for equality of productivity change at the aggregate level, which is very useful in practice (e.g., comparing productivity change of industries of an economy or groups of countries over time). Another potential avenue for future research is to improve the performance of the developed statistical inferences in small finite samples by correcting for the bias in the estimator of variance. This is analogous to what Simar and Zelenyuk (2020) derive for the context of efficiency scores, yet it is likely to require considerable elaboration and efforts for the more complex context of MPIs as our developments suggest, and so it is a subject in itself for future research endeavors.

Acknowledgments

The authors thank the editors and all anonymous referees for the fruitful feedback that helped improve this paper substantially. They also thank everyone who gave feedback on this paper, especially Bao Hoang Nguyen, Zhichao Wang, and Evelyn Smart as well as participants of various conferences and seminars at which this paper was presented, including the European Workshop on Efficiency and Productivity Analysis 2022. They also thank Zhichao Wang for technical assistance with data preparation and for checking the programming codes and computations. The Australian Government Research Training Program, University of Queensland, and Australian Research Council (whose support is acknowledged and appreciated) are not responsible for the views expressed in this paper.

Appendix A. Regularity Assumptions

This appendix includes the regularity assumptions needed to develop statistical properties for the MPI.

First, for each $t = 1, 2$, Ψ^t is assumed to satisfy common regularity assumptions in production theory (see, e.g., Färe and Primont 1995, Sickles and Zelenyuk 2019, for details), and in particular, the following.

Assumption A.1. *Ψ^t is closed and convex.*

Assumption A.2 (No Free Lunch). $(0, y) \notin \Psi^t \;\; \forall y \geq 0$.

Assumption A.3 (Strong Disposability of Inputs and Outputs). *If $(x, y) \in \Psi^t$, then $(x^*, y^*) \in \Psi^t$ for $x^* \geqq x$, $y^* \leqq y$.*

The next set of conditions are statistical in nature. The first is adapted from SZ2018 to ensure the key components of aggregate MPIs have (at least) the first two moments.

Assumption A.4. *The first two moments of $w^1 Y_i^1$ and $w^2 Y_i^2$ are finite for all $i = 1, \ldots, n$.*

The next set is analogous to those in KSW2021 as assumptions 2.4–2.7, 3.1, and 3.2, which are presented for the Farrell-type input-oriented efficiency measure. Here, we give their analogues for the output-oriented efficiency measure, which, to the best of our knowledge, have not been presented in the literature before, and we thank the anonymous referee for the nudge to do so.

$$\lambda(z|\Psi^t) = \lambda(x, y|\Psi^t) = \sup_{\lambda}\{\lambda : (x, \lambda y) \in \Psi^t\}. \tag{A.1}$$

Assumption A.5. (i) *The random vector (X_i^t, Y_i^t) possesses a joint density f^t with support $\mathcal{D}^t \subset \Psi^t$, and* (ii) *$f^t$ is continuously differentiable on $\mathcal{D}^t$.*

Assumption A.6. (i) *$\mathcal{D}^{t*} \subset \mathcal{D}^t$, where $\mathcal{D}^{t*} = \{(x, \lambda(x, y|\Psi^t)y) : (x, y) \in \mathcal{D}^t\}$,* (ii) *$\mathcal{D}^{t*}$ is compact, and* (iii) *$f^t(x, \lambda(x, y|\Psi^t)y) > 0$ for all $(x, y) \in \mathcal{D}^t$.*

Assumption A.7. *$\lambda(x,y|\Psi^t)$ is three times continuously differentiable on $\mathcal{D}^t$.*

Assumption A.8 (D^t Is Almost Strictly Convex). *For any $(x,y),(\tilde{x},\tilde{y}) \in \mathcal{D}^t$ with $\left(x, \frac{y}{\|y\|}\right) \neq \left(\tilde{x}, \frac{\tilde{y}}{\|\tilde{y}\|}\right)$, the set $\{(x^*,y^*):(x^*, y^*) = (1-\alpha)(x,y)+\alpha(\tilde{x},\tilde{y}), \alpha \in (0,1)\}$ is a subset of the interior of $\mathcal{D}^t$.*

Prior to mentioning Assumptions A.9 and A.10, we need additional definitions (analogous to KSW2021):

$$\mathcal{D}^t_{norm} = \left\{\left(\frac{x}{\|x\|}, \frac{y}{\|y\|}\right) : (x,y) \in \mathcal{D}^t\right\} \tag{A.2}$$

and

$$\tilde{g}_y\left(a\frac{x}{\|x\|}\right) = \max_{b>0}\left\{b : \left(a\frac{x}{\|x\|}, b\frac{y}{\|y\|}\right) \in \Psi^t\right\}. \tag{A.3}$$

Then, there exists some $\alpha^{x,y}_{max} > 0$ such that

$$\frac{\tilde{g}_y\left(\alpha^{x,y}_{max}\frac{x}{\|x\|}\right)}{\alpha^{x,y}_{max}} = \max_{a>0}\left\{\frac{\tilde{g}_y\left(a\frac{x}{\|x\|}\right)}{a} : \left(a\frac{x}{\|x\|}, \tilde{g}_y\left(a\frac{x}{\|x\|}\right)\frac{y}{\|y\|}\right) \in \Psi^t\right\}, \tag{A.4}$$

where $\alpha^{x,y}_{max} \in \mathbb{R}_+$ is necessarily uniquely defined if Ψ^t is strictly convex.

Assumption A.9. (i) *The support $\mathcal{D}^t \subset \Psi^t$ of f^t is such that, for any $\left(\frac{x}{\|x\|}, \frac{y}{\|y\|}\right) \in \mathcal{D}^t_{norm}$, we have $\left(\alpha^{x,y}_{max}\frac{x}{\|x\|}, \tilde{g}_y\left(\alpha^{x,y}_{max}\frac{x}{\|x\|}\right)\frac{y}{\|y\|}\right) \in \mathcal{D}^t$;* (ii) *there exists a $\delta > 0$ such that, for any $\left(\frac{x}{\|x\|}, \frac{y}{\|y\|}\right) \in \mathcal{D}^t_{norm}$, we also have $\left([\alpha^{x,y}_{max} - \delta]\frac{x}{\|x\|}, \tilde{g}_y\left([\alpha^{x,y}_{max} - \delta]\frac{x}{\|x\|}\right)\frac{y}{\|y\|}\right) \in \mathcal{D}^t$ and $\left([\alpha^{x,y}_{max} + \delta]\frac{x}{\|x\|}, \tilde{g}_y\left([\alpha^{x,y}_{max} + \delta]\frac{x}{\|x\|}\right)\frac{y}{\|y\|}\right) \in \mathcal{D}^t$; and* (iii) *there exists a constant $0 < M < \infty$ such that $\|y\| \leq M \ \forall (x,y) \in \mathcal{D}^t$.*

Assumption A.10. (i) *For $t \in \{1,2\}$, there are independent and identically distributed (iid) observations $(X^t_i, Y^t_i), i = 1,\ldots,n_t$ such that Assumptions A.1–A.9 are satisfied with respect to the underlying densities f^t with supports $\mathcal{D}^t$;* (ii) *$\mathcal{D}^1_{norm} = \mathcal{D}^2_{norm}$;* (iii) *for some $n \leq \min\{n_1, n_2\}$, the observations $((X^1_i, Y^1_i), (X^2_i, Y^2_i)), i = 1,\ldots,n$ are iid and their joint distribution posseses a continuous density f_{12} with support $\mathcal{D}^1 \times \mathcal{D}^2$;* (iv) *for any $i = 1,\ldots,n_1$, (X^1_i, Y^1_i) is independent of (X^2_j, Y^2_j) for all $j = 1,\ldots,n_2$ with $i \neq j$;* (v) *for any $i = 1,\ldots,n_2$, (X^2_i, Y^2_i) is independent of (X^1_j, Y^1_j) for all $j = 1,\ldots,n_1$ with $i \neq j$.*

The meaning and importance of these assumptions is analogous to what is explained in fair detail in KSW2021, and hence, here we provide only a brief intuition for the last two assumptions that are the most technical, very important, and less common in the related literature. Assumption A.9 basically ensures that (i) the output-oriented distance to the frontier of the conical hull can be estimated consistently with the observed data, (ii) without running into a boundary problem, and (iii) its required moments exist. Meanwhile, Assumption A.10 regularizes the dynamic framework by guaranteeing that the DEA estimators of the output-oriented and hyperbolic distances to the frontier of the conical hull follow the asymptotic distributions analogous to those in theorems 3.1 and 3.2 of KSW2021 and ensures that the dynamic versions of these estimators are comparable as well as being asymptotically well-defined and possess the same convergence rates as the standard DEA estimators. It also requires that the observations are independent across different DMUs (although note that a dependence of observations for the same DMU across time is allowed).

Appendix B. Asymptotic Theory When the True Efficiency Is Known

In this appendix, we develop central limit theorems in relation to $\overline{M}$ by assuming that the true efficiency, $\lambda_C(\cdot|\Psi^1)$ and $\lambda_C(\cdot|\Psi^2)$, are known. Deriving asymptotic theory under this assumption is important because it enlightens the statistical essence of the aggregation form $\overline{M}$ and helps identify the underlying parameters that these estimators gauge. Moreover, the results derived here also provide a statistical grounding for the main sections of the paper in which we develop an asymptotic theory in the absence of the knowledge of $\lambda_C(\cdot|\Psi^1)$ and $\lambda_C(\cdot|\Psi^2)$ as usually happens in reality. Although we focus on the DEA estimator in this paper, the theory presented in this appendix is equally useful for alternative estimators, including stochastic frontier analyses (SFA).

Given the notation and definitions outlined in Section 4, consider iid random variables

$$T_i = [U_{1,i}, U_{2,i}, U_{3,i}, U_{4,i}, U_{5,i}, U_{6,i}]' \quad (i = 1,\ldots,n) \tag{B.1}$$

and denote their means and variances by

$$\mu = E(T_i) = [\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6]', \Sigma = Var(T_i) = [\sigma_{jk}]_{j,k\in\{1,2,3,4,5,6\}}. \tag{B.2}$$

By the standard central limit theorem (see, e.g., Van der Vaart 2000), we have

$$\sqrt{n}(\hat{\mu}_n - \mu) \xrightarrow{d} \mathcal{N}(0, \Sigma). \tag{B.3}$$

Here, $\hat{\mu}_n = [\hat{\mu}_{1,n}, \hat{\mu}_{2,n}, \hat{\mu}_{3,n}, \hat{\mu}_{4,n}, \hat{\mu}_{5,n}, \hat{\mu}_{6,n}]'$.

Define the function $\phi : (0,\infty)^6 \to \mathbb{R}^1$ as

$$\phi(\eta_1, \eta_2, \eta_3, \eta_4, \eta_5, \eta_6) = \frac{1}{2}(\log\eta_3 + \log\eta_4 - \log\eta_1 - \log\eta_2) + \log\eta_5 - \log\eta_6. \tag{B.4}$$

Then, under standard regularity conditions, we can apply the delta method (see, e.g., theorem 3.1 of Van der Vaart 2000) to (B.3) and obtain

$$\sqrt{n}(\phi(\hat{\mu}_n) - \phi(\mu)) \xrightarrow{d} \mathcal{N}(0, [\nabla\phi(\mu)]'\Sigma[\nabla\phi(\mu)]), \tag{B.5}$$

Or, equivalently,

$$\sqrt{n}(\hat{\xi}_n - \xi) \xrightarrow{d} \mathcal{N}(0, V_\xi), \tag{B.6}$$

where

$$V_\xi = nVAR(\hat{\xi}_n) = [\nabla\phi(\mu)]'\Sigma[\nabla\phi(\mu)], \tag{B.7}$$

and $\nabla\phi(\cdot)$ is the column vector of the gradient of $\phi(\cdot)$. Formally, $\nabla\phi(\mu) = \left[\frac{\partial\phi}{\partial\eta_j}(\mu)\right]$, where

$$\frac{\partial\phi}{\partial\eta_1}(\mu) = -\frac{1}{2\mu_1}, \quad \frac{\partial\phi}{\partial\eta_2}(\mu) = -\frac{1}{2\mu_2}, \quad \frac{\partial\phi}{\partial\eta_3}(\mu) = \frac{1}{2\mu_3},$$
$$\frac{\partial\phi}{\partial\eta_4}(\mu) = \frac{1}{2\mu_4}, \quad \frac{\partial\phi}{\partial\eta_5}(\mu) = \frac{1}{\mu_5}, \quad \frac{\partial\phi}{\partial\eta_6}(\mu) = -\frac{1}{\mu_6}. \tag{B.8}$$

Evidently, $U_{5,i}$ and $U_{6,i}$ are observable for all $i = 1,\ldots,n$. Under the assumption that the functions $\lambda_C(\cdot|\Psi^1)$ and $\lambda_C(\cdot|\Psi^2)$ are known, $U_{s,i}$ $(s = 1,\ldots,4; i = 1,\ldots,n)$ are also observable, and so is $\widehat{\xi}_n$. Let $\widehat{V}_{\xi,n}$ denote the empirical version of V_ξ, where μ_s is replaced by $\widehat{\mu}_{s,n}$ and σ_{jk} is replaced by $\widehat{\sigma}_{jk,n}$ with

$$\widehat{\sigma}_{jk,n} = n^{-1}\sum_{i=1}^{n}(U_{j,i} - \widehat{\mu}_{j,n})(U_{k,i} - \widehat{\mu}_{k,n}), \quad j,k = 1,2,\ldots,6. \tag{B.9}$$

It is well-known that $\widehat{\mu}_{s,n} \xrightarrow{p} \mu_s$ and $\widehat{\sigma}_{jk,n} \xrightarrow{p} \sigma_{jk}$ $(s,j,k = 1,\ldots,6)$ as $n \to \infty$. Hence, by Slutsky's theorem and the continuous mapping theorem (see, e.g., Van der Vaart 2000), we have $\widehat{V}_{\xi,n} \xrightarrow{p} V_\xi$. Combining these results with (B.6) and Slutsky's theorem, we can obtain

$$\sqrt{n}\left(\frac{\widehat{\xi}_n - \xi}{\sqrt{\widehat{V}_{\xi,n}}}\right) \xrightarrow{d} \mathcal{N}(0,1), \tag{B.10}$$

which, in turn, can be used to make inferences about ξ. In particular, when the true efficiency is known, the asymptotic $100(1-\alpha)\%$ symmetric confidence interval for ξ is given by

$$\left[\widehat{\xi}_n \pm \Phi^{-1}_{1-\alpha/2}\sqrt{\widehat{V}_{\xi,n}/n}\right], \tag{B.11}$$

where $\Phi^{-1}_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of the standard normal distribution.

Of course, in reality, $\lambda_C(\cdot|\Psi^1)$ and $\lambda_C(\cdot|\Psi^2)$ are unknown and need to be estimated. In general, one has to check whether the replacements of truth with any estimates (especially if they are biased) preserves the validity of such theorems. In the main text of the paper, we show that this is not the case for the DEA estimators because of a slow bias that needs to be corrected, and hence, the new central limit theorems need to be derived, which is the goal of the paper. Although the derivation of these theorems is much more involved, the simpler theory in this appendix is still novel and serves as a useful guide for such derivations as well as, potentially, for the derivations involving other estimators, such as SFA, etc.

Endnotes

[1] For example, a Google Scholar web search found about 24,400 results for "Malmquist productivity index" as of March 11, 2019, and 32,200 as of September 27, 2021.

[2] For applications of the MPI, see, for example, Färe et al. (1994), Johnson and Ruggiero (2014), Brennan et al. (2014), and Simar and Wilson (2019) as well as an overview in Sickles and Zelenyuk (2019) and references therein.

[3] See large commercial bank releases for March 31, 2019 (data file). Retrieved June 21, 2019, from https://www.federalreserve.gov/releases/lbr/.

[4] A similar theory can also be developed for other indices based on the methodology introduced in this paper.

[5] For a related discussion, see Zelenyuk (2014), who call it a conical closure of Ψ^t and consider its scale-homothetic decompositions.

[6] For practical reasons, we assume away the so-called singularity cases (Sickles and Zelenyuk 2019), in which the efficiency measures are either zero or ∞, to make sure the MPI is well-defined.

[7] This aggregation appears to be more appropriate than the conventional arithmetic and geometric means as the weight β_i^s represents the output share of DMU i in the period s and, hence, reflects its relative economic importance in the aggregate indices. As pointed out by Zelenyuk (2006), these weights are "not *ad hoc* but are derived from economic principles (agents' optimization behavior)" and using $\overline{M}$ enables the group revenue analog of the MPI to be decomposed in the same way as for the individual revenue analog of the MPI. Empirical applications of the system of weights β_i^s as well as the aggregation $\overline{M}$ can be found in a number of studies (e.g., Pilyavsky and Staat 2008, Gitto and Mancuso 2015) although the asymptotic theory for a well-grounded statistical analysis has not been developed yet, which is the goal of this study. (For more details of this measure, see Zelenyuk 2006).

[8] Other popular approaches to measuring performance include SFA, including its nonparametric versions, as well as stochastic DEA (e.g., see Olesen and Petersen 2016, Parmeter and Zelenyuk 2019 for recent reviews and references therein).

[9] For $a, b \in \mathbb{R}^m$, "$a \geqq b$" or "$b \leqq a$" means $a - b \in \mathbb{R}^m_+$, "$a \geq b$" or "$b \leq a$" means $a - b \in \mathbb{R}^m_+ \backslash \{0_m\}$, "$a > b$" or "$b < a$" means $a - b \in \mathbb{R}^m_{++}$.

[10] Theorem 1 in this paper is an analogue of theorem 3.1 of KSW2015, corollary 1 of SZ2018, and theorem 3.4 of KSW2021.

[11] Theorem 3 here is an analogue of theorem 4.1 of KSW2015 and theorem 1 of SZ2018 but encompasses an additional case (part (iv)), which corresponds to the covariance between random variables containing efficiency.

[12] It is worth noting that the use of the uniform delta method has recently received attention in other areas of the econometrics literature (e.g., see Kasy 2019). Our use of this method is different and, to the best of our knowledge, is novel to the productivity literature and, perhaps, to the econometrics literature in general. In particular, Kasy (2019) focuses on conditions for the remainder to converge uniformly. Meanwhile, our paper focuses on the rates of convergence of the remainders.

[13] Note that $\widetilde{n}$ depends on the original sample n and κ (and κ depends on the type of the estimator), but to simplify the notation and avoid possible confusion, we drop κ from the subscript. We thank the anonymous referee for encouraging this simplification.

[14] Note that, when $\kappa = 1/2$, Theorems 5 and 6 are equivalent.

[15] Following the literature, we set $L = 10$ in our Monte Carlo simulations. A sensitivity check for different choices of L confirms the robustness of the results.

[16] In other words, the recentered confidence intervals overcover when n is sufficiently large.

[17] The definition of developed versus developing is similar to Henderson and Zelenyuk (2007) except that here the group of developing countries is larger.

References

Badunenko O, Henderson DJ, Zelenyuk V (2008) Technological change and transition: Relative contributions to worldwide growth during the 1990s. *Oxford Bull. Econom. Statist.* 70(4):461–492.

Brennan S, Haelermans C, Ruggiero J (2014) Nonparametric estimation of education productivity incorporating nondiscretionary inputs with an application to Dutch schools. *Eur. J. Oper. Res.* 234(3):809–818.

Caves DW, Christensen LR, Diewert WE (1982) The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica* 50(6):1393–1414.

Ebert U, Welsch H (2004) Meaningful environmental indices: A social choice approach. *J. Environ. Econom. Management* 47(2):270–283.

Färe R, Grosskopf S (2004) *New Directions: Efficiency and Productivity* (Springer).

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

Färe R, Zelenyuk V (2003) On aggregate Farrell efficiencies. *Eur. J. Oper. Res.* 146(3):615–620.

Färe R, Grosskopf S, Lovell CK (1985) *The Measurement of Efficiency of Production*, vol. 6 (Springer, New York).

Färe R, Grosskopf S, Norris M, Zhang Z (1994) Productivity growth, technical progress, and efficiency change in industrialized countries. *Amer. Econom. Rev.* 84(1):66–83.

Färe R, He X, Li S, Zelenyuk V (2019) A unifying framework for Farrell profit efficiency measurement. *Oper. Res.* 67(1):183–197.

Farrell MJ (1957) The measurement of productive efficiency. *J. Roy. Statist. Soc. Ser. A* 120(3):253–290.

Gitto S, Mancuso P (2015) The contribution of physical and human capital accumulation to Italian regional growth: A nonparametric perspective. *J. Production Anal.* 43(1):1–12.

Grifell-Tatjé E, Lovell CK (1995) A note on the Malmquist productivity index. *Econom. Lett.* 47(2):169–175.

Henderson DJ, Zelenyuk V (2007) Testing for (efficiency) catching-up. *Southern Econom. J.* 73(4):1003–1019.

Johnson AL, Ruggiero J (2014) Nonparametric measurement of productivity and efficiency in education. *Ann. Oper. Res.* 221(1):197–210.

Kasy M (2019) Uniformity and the delta method. *J. Econom. Methods* 8(1):1–19.

Kneip A, Simar L, Wilson PW (2015) When bias kills the variance: Central limit theorems for DEA and FDH efficiency scores. *Econometric Theory* 31(2):394–422.

Kneip A, Simar L, Wilson PW (2016) Testing hypotheses in nonparametric models of production. *J. Bus. Econom. Statist.* 34(3):435–456.

Kneip A, Simar L, Wilson PW (2021) Inference in dynamic, nonparametric models of production: Central limit theorems for Malmquist indices. *Econometric Theory* 37(3):537–572.

Kumar S, Russell RR (2002) Technological change, technological catch-up, and capital deepening: Relative contributions to growth and convergence. *Amer. Econom. Rev.* 92(3):527–548.

Olesen OB, Petersen NC (2016) Stochastic data envelopment analysis—A review. *Eur. J. Oper. Res.* 251(1):2–21.

Parmeter CF, Zelenyuk V (2019) Combining the virtues of stochastic frontier and data envelopment analysis. *Oper. Res.* 67(6):1628–1658.

Pastor JM, Perez F, Quesada J (1997) Efficiency analysis in banking firms: An international comparison. *Eur. J. Oper. Res.* 98(2):395–407.

Pilyavsky A, Staat M (2008) Efficiency and productivity change in Ukrainian healthcare. *J. Production Anal.* 29(2):143–154.

Ray SC, Desli E (1997) Productivity growth, technical progress, and efficiency change in industrialized countries. *Amer. Econom. Rev.* 87(5):1033–1039.

Sickles R, Zelenyuk V (2019) *Measurement of Productivity and Efficiency: Theory and Practice* (Cambridge University Press, New York).

Simar L, Wilson PW (2019) Central limit theorems and inference for sources of productivity change measured by nonparametric Malmquist indices. *Eur. J. Oper. Res.* 277(2):756–769.

Simar L, Zelenyuk V (2018) Central limit theorems for aggregate efficiency. *Oper. Res.* 66(1):137–149.

Simar L, Zelenyuk V (2020) Improving finite sample approximation by central limit theorems for estimates from data envelopment analysis. *Eur. J. Oper. Res.* 284(3):1002–1015.

ten Raa T (2011) Benchmarking and industry performance. *J. Production Anal.* 36(3):285–292.

Tortosa-Ausina E, Grifell-Tatjé E, Armero C, Conesa D (2008) Sensitivity analysis of efficiency and Malmquist productivity indices: An application to Spanish savings banks. *Eur. J. Oper. Res.* 184(3):1062–1084.

van der Vaart AW (2000) *Asymptotic Statistics*, vol. 3 (Cambridge University Press, New York).

Walheer B (2018) Aggregation of metafrontier technology gap ratios: The case of European sectors in 1995–2015. *Eur. J. Oper. Res.* 269(3):1013–1026.

Walheer B (2019) Aggregating Farrell efficiencies with private and public inputs. *Eur. J. Oper. Res.* 276(3):1170–1177.

Wang H, Ang B, Wang Q, Zhou P (2017) Measuring energy performance with sectoral heterogeneity: A non-parametric frontier approach. *Energy Econom.* 62:70–78.

Ylvinger S (2000) Industry performance and structural efficiency measures: Solutions to problems in firm models. *Eur. J. Oper. Res.* 121(1):164–174.

Zelenyuk V (2006) Aggregation of Malmquist productivity indexes. *Eur. J. Oper. Res.* 174(2):1076–1086.

Zelenyuk V (2014) Scale efficiency and homotheticity: Equivalence of primal and dual measures. *J. Production Anal.* 42(1):15–24.

Manh Pham obtained his PhD in econometrics at the University of Queensland in 2019, receiving Faculty Excellence in Research Award for Higher Degree Research. He is currently a senior quantitative analyst in Australia.

Léopold Simar is an emeritus professor at the Institute of Statistics, Biostatistics and Actuarial Sciences of the Université Catholique de Louvain (UCL). He founded the Institute of Statistics at UCL in 1992 and chaired the Institute from its creation till 2004. His research is in the fields of mathematical statistics and econometrics methods, including the nonparametric approaches, bootstrap, etc. In particular, he is recognized as an expert on frontier estimation and efficiency analysis.

Valentin Zelenyuk is a professor in econometrics at the University of Queensland, where he also served as the Director of Centre for Efficiency and Productivity Analysis. His research focuses on three directions: (i) economic theory of production, productivity, and efficiency and related aggregation issues; (ii) econometric/statistical estimation (mainly nonparametric and semiparametric); and (iii) econometric applications (mainly to healthcare, banking, and cross-countries analysis).