STATISTICAL INFERENCE FOR THE AGGREGATE SOURCES OF PRODUCTIVITY CHANGE MEASURED BY MALMQUIST PRODUCTIVITY INDICES

Léopold Simar, Valentin Zelenyuk, Shirong Zhao

LIDAM Discussion Paper ISBA
2026 / 24

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

Statistical Inference for the Aggregate Sources of Productivity Change Measured by Malmquist Productivity Indices

Léopold Simar* Valentin Zelenyuk† Shirong Zhao‡

June 13, 2026

Abstract

The Malmquist productivity index (MPI) is one of the most widely used tools for measuring productivity change over time. A recent contribution by Pham et al. (2024) develops statistical inference procedures for the aggregate (or weighted) MPI of a group of firms. Building on this framework, the present paper develops statistical inference for the aggregate sources of productivity change measured by MPI. In particular, we derive the central limit theorems of the aggregated components, thereby enabling hypothesis testing for changes in technology, efficiency, and related sources. We illustrate the developed theoretical results by analyzing the sources of productivity change for 84 countries. Moreover, analogous theoretical results can be established for alternative productivity measures and their decompositions, including the Hicks–Moorsteen productivity index, the Luenberger productivity index, the Malmquist–Luenberger productivity index, etc.

Keywords: Data Envelopment Analysis, Malmquist Productivity Index, Aggregation, Central Limit Theorem, Efficiency Change, Technology Change

JEL Classification: C12,C14,C43,C67

*Simar: Institut de Statistique, Biostatistique et Sciences Actuarielles (ISBA), LIDAM, UCLouvain, 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 (Corresponding Author): School of Finance, Dongbei University of Finance and Economics, Dalian, Liaoning 116025, China; email: shironz@163.com.

1 Introduction

The Malmquist productivity index (MPI) proposed by Caves et al. (1982) is widely recognized as a benchmark tool for measuring intertemporal productivity change.[1] As of May 7th, 2026, a Google search for the keyword “Malmquist productivity index” returns approximately $52,200$ results, illustrating the widespread use and continuing popularity of the MPI in the literature. Representative applications of the MPI include Färe et al. (1994b), Kumar and Russell (2002), Henderson and Russell (2005), Badunenko and Romero-Ávila (2013), Ramakrishna et al. (2016), Pastor et al. (2020), Wilson and Zhao (2025), Villa and Lozano (2025), and Ma et al. (2026), among others.[2]

The MPI admits a variety of decompositions that allow researchers to identify the underlying sources of productivity change, such as technological change, efficiency change, scale efficiency change and other factors (Färe et al., 1992; Färe et al., 1994a; Wheelock and Wilson, 1999). Recently, Kneip et al. (2021) established a unified statistical framework for MPI-based productivity change. Simar and Wilson (2019) developed the corresponding theoretical results for its decompositions, which equip researchers with formal tools to conduct statistical inference—for example, to test whether efficiency change is statistically significant. An example of empirical application of this inferential framework is provided in Simar and Wilson (2023), who analyze productivity change and its determinants at the country level in a setting analogous to Färe et al. (1994b), while explicitly incorporating statistical inference into the decomposition analysis. However, Simar and Wilson (2019) focus exclusively on the unweighted (or simple mean) approach, treating each individual decision-making unit equally in the aggregation process.

A more recent contribution by Pham et al. (2024) advances MPI literature by developing formal statistical inference for the aggregate (weighted) MPI computed across a group of decision-making units, building upon and complementing the unweighted framework of Kneip et al. (2021). In this setting, the aggregation weights are grounded in economic theory: they correspond to revenue shares when the MPI is defined in an output orientation and to cost

[1] Another well-known productivity measure is the Hicks–Moorsteen productivity index, introduced by Diewert (1992) and Bjurek (1996), with its statistical properties later developed by Simar et al. (2025b).

[2] For a comprehensive review of productivity measurement and related methodologies, see Sickles and Zelenyuk (2019).

shares when it is defined in an input orientation.[3] Importantly, Pham et al. (2024) did not establish the corresponding statistical theory for the aggregate sources of MPI-based productivity change, such as efficiency change and technological change—components that are often central to economic interpretation and policy analysis. Statistical inference for these components therefore remains an open question, and we try to address it in this work.

It turns out that estimates of MPI and its sources—and the resulting policy implications—may differ substantially depending on the weighting scheme used for aggregation. Table 1 presents a numerical example involving six firms, each using two inputs to produce one output, to illustrate the differences between weighted and unweighted MPI and its components. When output weights (or revenue shares) are used to aggregate the output-oriented MPI, the aggregate estimates for technological change, efficiency change, and MPI, are 1.02, 0.93, and 0.95, respectively. In contrast, the corresponding simple mean estimates are 1.03, 1.03, and 1.06. Thus, the weighted aggregation suggests a decline in productivity of about 5%, driven by a 2% increase due to technological change and a 7% decrease due to efficiency change. On the other hand, equal weighting implies a very different picture: an increase in productivity of approximately 6%, where 3% increase is due to technological change and 3% increase due to efficiency change.

This example illustrates that weighted and unweighted methods may yield substantially different result—not only in the direction of productivity change, but also in the attribution of the underlying sources of that change. Importantly, such differences can translate into materially different policy conclusions. The example therefore highlights the importance of weighted aggregation, where weights reflect the relative economic significance of firms within a group, as a complement to simple averaging.

The present paper addresses this gap by deriving the asymptotic results for the weighted (aggregate) sources of MPI, thereby complementing the theoretical results for unweighted decompositions developed by Simar and Wilson (2019). Importantly, extending the analysis from unweighted to weighted sources is nontrivial. The weighted decomposition involves additional terms and a more intricate dependence structure, which substantially increases the technical complexity relative to the unweighted case.

[3] For a more detailed discussion, see Zelenyuk (2022, 2023, 2024) and the review chapter in Sickles and Zelenyuk (2019, Chapter 5).

Table 1: A Numerical Example

	Period 1			Period 2			Technology	Efficiency	MPI
Firm	Input 1	Input 2	Output	Input 1	Input 2	Output	change	change	
1	1.00	1.00	1.00	1.00	1.00	2.00	0.96	2.08	2.00
2	1.00	2.00	2.00	1.00	2.00	2.00	1.07	0.94	1.00
3	1.00	3.00	3.00	1.00	3.00	4.00	1.33	1.00	1.33
4	2.00	1.00	4.00	2.00	1.00	4.00	1.00	1.00	1.00
5	2.00	2.00	5.00	2.00	2.00	4.00	0.96	0.83	0.80
6	2.00	3.00	6.00	2.00	3.00	4.00	0.93	0.71	0.67

	Technology change	Efficiency change	MPI
Weighted	1.02	0.93	0.95
Unweighted	1.03	1.03	1.06

NOTE: The output price is assumed to be unity for all firms across both periods. The computations were done using the conical efficiency estimator as described in (3.2) in Section 3.

In particular, we derive central limit theorems for the aggregate sources of productivity change measured by MPI. These results provide a rigorous foundation for conducting hypothesis tests and constructing confidence intervals for technological change, efficiency change, and other components. Consequently, the proposed framework contributes to the literature by providing statistical inference on the drivers underlying the changes in aggregate productivity.

To illustrate the empirical applicability of the theory, we implement the proposed methods using data from the Penn World Table to examine productivity dynamics for a panel of 84 countries. The empirical analysis reveals substantial differences between bias-corrected and uncorrected estimates for both the aggregate MPI and its decomposed components, underscoring the practical importance of bias adjustment. Moreover, while both weighted and unweighted approaches suggest that aggregate productivity increased by approximately 50% between 1990 and 2023, they yield markedly different interpretations of the sources of this growth. The unweighted measure attributes productivity gains to improvements in both conical technical efficiency and technological progress. In contrast, the weighted aggregate measure indicates that productivity growth is predominantly driven by gains in conical technical efficiency, with technological progress playing no statistically significant role. Taken together, our findings demonstrate that the choice between weighted and unweighted aggregation substantively affects economic interpretation. Examining both measures is therefore essential for obtaining a comprehensive and reliable assessment of productivity change and its underlying determinants within a group.

The remainder of this study is organized as follows. We present the theoretical foundation, covering production economics, the aggregate MPI and its components in Section 2. Section 3 introduces the estimators obtained using Data Envelopment Analysis. Section 4 establishes the statistical results for aggregate sources of productivity change measured by MPI. Section 5 explores one special case of production technology that needs to be taken care of. In Section 6, we illustrate the developed statistical results for aggregate sources of MPI by using one real data set. Section 7 concludes and recommends the future research directions.

2 The Theoretical Background

We first introduce the theoretical foundation, covering production economics, the aggregate MPI and its components. The notation used below generally follows Simar and Wilson (2019) and Simar et al. (2025b).

2.1 The Production Economics Model

Let $x \in \mathbb{R}_+^p$ be the column vector of inputs, and $y \in \mathbb{R}_+^q$ be the column vector of outputs. The production technology corresponding to time t is described by the technology set:

$$\mathcal{P}^t = \{(x, y) \mid y \text{ can be produced from } x \text{ at time } t\}, \tag{2.1}$$

which defines all feasible input-output combinations attainable with the available technology at time t.[4]

The *production frontier*, denoted as $\mathcal{P}^{t\partial}$, characterizes the production boundary of the $\mathcal{P}^t$ can be defined as:

$$\mathcal{P}^{t\partial} := \left\{(x, y) \in \mathcal{P}^t \mid (x, \lambda y) \notin \mathcal{P}^t \text{ for any } \lambda > 1\right\}, \tag{2.2}$$

indicating those production points that are efficient and cannot be proportionally scaled further in output without violating feasibility while keeping inputs and technology fixed.

[4] We assume standard regularity conditions for $\mathcal{P}^t$, following Appendix A in Pham et al. (2024).

Technical efficiency for a given point (x, y) measures the distance to the frontier $\mathcal{P}^{t\partial}$. While various efficiency measures exist, we focus on the output-oriented approach (Farrell, 1957), which is defined as:

$$\lambda(x, y \mid \mathcal{P}^t) := \sup\left\{\lambda > 0 \mid (x, \lambda y) \in \mathcal{P}^t\right\}, \tag{2.3}$$

which captures the maximum factor by which all outputs can be proportionally expanded while keeping all inputs and $\mathcal{P}^t$ fixed.

The *conical closure* of $\mathcal{P}^t$ can be denoted as:

$$\mathcal{C}(\mathcal{P}^t) := \left\{(\tilde{x}, \tilde{y}) \mid \tilde{x} = \rho x,\ \tilde{y} = \rho y,\ \forall \rho \in \mathbb{R}_+,\ (x, y) \in \mathcal{P}^t\right\}. \tag{2.4}$$

The efficiency measured in output orientation relative to $\mathcal{C}(\mathcal{P}^t)$ (i.e., the conical technical efficiency) is given by:

$$\lambda_C(x, y \mid \mathcal{P}^t) := \lambda(x, y \mid \mathcal{C}(\mathcal{P}^t)) = \sup\left\{\lambda > 0 \mid (x, \lambda y) \in \mathcal{C}(\mathcal{P}^t)\right\}. \tag{2.5}$$

Under constant returns to scale (CRS) for $\mathcal{P}^t$, we have $\mathcal{P}^t = \mathcal{C}(\mathcal{P}^t)$ and in this case (2.3) and (2.5) coincide; otherwise, $\mathcal{P}^t \subset \mathcal{C}(\mathcal{P}^t)$.

2.2 The Malmquist Productivity Indices

Suppose we collect a random sample of $\mathcal{X}_n = \{(X_i^1, Y_i^1), (X_i^2, Y_i^2)\}_{i=1}^n$, which contains the input–output pairs (X_i^t, Y_i^t) among n firms in time $t \in \{1, 2\}$. We further denote $Z_i^t = (X_i^t, Y_i^t)$ and $\mathcal{X}_n^t = \{Z_i^t\}_{i=1}^n$ for time t. The productivity change for the ith firm from time 1 to time 2, measured by the MPI (Caves et al., 1982), is given by

$$\mathcal{M}_i := \left(\frac{\lambda_C(Z_i^2 \mid \mathcal{P}^1)}{\lambda_C(Z_i^1 \mid \mathcal{P}^1)} \times \frac{\lambda_C(Z_i^2 \mid \mathcal{P}^2)}{\lambda_C(Z_i^1 \mid \mathcal{P}^2)}\right)^{-1/2}, \tag{2.6}$$

where, $\mathcal{M}_i > 1, = 1,$ or < 1, reflects whether the productivity of firm i has improved, stayed the same, or declined between time 1 and time 2.[5]

[5] It is worth emphasizing that the MPI is measured relative to the conical technology frontier. We do not impose CRS for the technology, and the technology may exhibit either VRS or CRS.

The simple (or equally-weighted) geometric mean MPI for n firms is given by

$$\overline{\mathbb{M}} = \left(\prod_{i=1}^{n} \mathcal{M}_i \right)^{1/n}, \tag{2.7}$$

which is commonly used in the literature. However, this measure assigns equal weight to each firm. As discussed in the Introduction, aggregation is not neutral: ignoring economically meaningful weights may lead to systematically different conclusions regarding the sources of productivity growth and the level of productivity growth itself.

An alternative approach is to incorporate the economic significance of each firm—such as its revenue—by constructing a revenue-weighted aggregate MPI (Zelenyuk, 2006; Pham et al., 2024), defined as

$$\overline{\mathcal{M}} := \left(\frac{\sum_{i=1}^{n} S_i^2 \lambda_C(Z_i^2 \mid \mathcal{P}^1)}{\sum_{i=1}^{n} S_i^1 \lambda_C(Z_i^1 \mid \mathcal{P}^1)} \times \frac{\sum_{i=1}^{n} S_i^2 \lambda_C(Z_i^2 \mid \mathcal{P}^2)}{\sum_{i=1}^{n} S_i^1 \lambda_C(Z_i^1 \mid \mathcal{P}^2)} \right)^{-1/2}, \tag{2.8}$$

where

$$S_i^t = \frac{w_y^t Y_i^t}{\sum_{i=1}^{n} w_y^t Y_i^t}, \tag{2.9}$$

is the weight representing firm i's share in time t, and $w_y^t \in \mathbb{R}_{++}^q$ is the row vector of output prices, which is assumed to be uniform across all firms in that time.[6] These weights reflect economic valuation through prices, thereby ensuring that aggregate productivity measures are consistent with economic aggregation principles.[7]

2.3 The Decomposition for the Aggregate MPI

The literature has proposed various decompositions of MPI to examine the sources of any changes in productivity. Färe et al. (1992) introduced a decomposition of the individual MPI that separates it into two distinct sources: conical efficiency change and technological change. This decomposition was later extended to the aggregate MPI by Zelenyuk (2006). The decomposition proposed by Zelenyuk (2006) for the aggregate MPI defined in (2.8) is

[6] To be well-defined, all the components in (2.6) and (2.8) are assumed to be non-zero.

[7] It is worth noting that the statistical results developed in Section 4 remain valid even if output prices differ across firms.

given by

$$\overline{\mathcal{M}} := \underbrace{\left(\frac{\sum_{i=1}^n S_i^2 \lambda_C(Z_i^2 \mid \mathcal{P}^2)}{\sum_{i=1}^n S_i^1 \lambda_C(Z_i^1 \mid \mathcal{P}^1)} \right)^{-1}}_{\overline{\mathcal{E}}_1} \times \underbrace{\left(\frac{\sum_{i=1}^n S_i^2 \lambda_C(Z_i^2 \mid \mathcal{P}^1)}{\sum_{i=1}^n S_i^2 \lambda_C(Z_i^2 \mid \mathcal{P}^2)} \times \frac{\sum_{i=1}^n S_i^1 \lambda_C(Z_i^1 \mid \mathcal{P}^1)}{\sum_{i=1}^n S_i^1 \lambda_C(Z_i^1 \mid \mathcal{P}^2)} \right)^{-1/2}}_{\overline{\mathcal{T}}_1}, \tag{2.10}$$

where $\overline{\mathcal{E}}_1$ and $\overline{\mathcal{T}}_1$ denote the conical efficiency change (movements toward or away from the conical frontier) and the conical technological change (shifts in the conical production frontier) measured based on the aggregate method. More specifically, $\overline{\mathcal{E}}_1$ measures whether firms are closer to the conical production frontier in time 2 than they were in time 1, with distances evaluated relative to the contemporaneous conical frontiers in each time. In turn, $\overline{\mathcal{T}}_1$ captures the shift in the conical production frontier itself, indicating whether the conical frontier has improved from time 1 to time 2. We have $\overline{\mathcal{E}}_1 \geq 1$ if and only if conical technical efficiency increases from time 1 to time 2. Similarly, $\overline{\mathcal{T}}_1 \geq 1$ if and only if the conical production frontier improves from time 1 to time 2.

Färe et al. (1994a) propose an alternative decomposition for the individual MPI. By adapting the ideas from Zelenyuk (2006), its aggregate analogue is

$$\overline{\mathcal{M}} := \underbrace{\left(\frac{\sum_{i=1}^n S_i^2 \lambda(Z_i^2 \mid \mathcal{P}^2)}{\sum_{i=1}^n S_i^1 \lambda(Z_i^1 \mid \mathcal{P}^1)} \right)^{-1}}_{\overline{\mathcal{E}}_2} \times \underbrace{\left(\frac{\left[\sum_{i=1}^n S_i^2 \lambda_C(Z_i^2 \mid \mathcal{P}^2)\right] / \left[\sum_{i=1}^n S_i^2 \lambda(Z_i^2 \mid \mathcal{P}^2)\right]}{\left[\sum_{i=1}^n S_i^1 \lambda_C(Z_i^1 \mid \mathcal{P}^1)\right] / \left[\sum_{i=1}^n S_i^1 \lambda(Z_i^1 \mid \mathcal{P}^1)\right]} \right)^{-1}}_{\overline{\mathcal{S}}_1} \times \overline{\mathcal{T}}_1, \tag{2.11}$$

where $\overline{\mathcal{S}}_1$ quantifies aggregate scale efficiency change, $\overline{\mathcal{E}}_2$ quantifies aggregate pure technical efficiency change, and $\overline{\mathcal{E}}_1 = \overline{\mathcal{E}}_2 \times \overline{\mathcal{S}}_1$. $\overline{\mathcal{S}}_1 \geq 1$ if and only if there is an improvement in scale efficiency between period 1 and period 2, and similarly $\overline{\mathcal{E}}_2 \geq 1$ corresponds to an increase in pure technical efficiency.

Ray and Desli (1997) propose another decomposition for the individual MPI; by adapting the ideas from Zelenyuk (2006), its aggregate analogue can be defined as

$$\begin{aligned}\overline{\mathcal{M}} := \overline{\mathcal{E}}_2 \times \underbrace{\left(\frac{\sum_{i=1}^n S_i^2 \lambda(Z_i^2 \mid \mathcal{P}^1)}{\sum_{i=1}^n S_i^2 \lambda(Z_i^2 \mid \mathcal{P}^2)} \times \frac{\sum_{i=1}^n S_i^1 \lambda(Z_i^1 \mid \mathcal{P}^1)}{\sum_{i=1}^n S_i^1 \lambda(Z_i^1 \mid \mathcal{P}^2)}\right)^{-1/2}}_{\overline{\mathcal{T}}_2} \times \\ \underbrace{\left(\frac{\left[\sum_{i=1}^n S_i^2 \lambda_C(Z_i^2 \mid \mathcal{P}^1)\right]/\left[\sum_{i=1}^n S_i^2 \lambda(Z_i^2 \mid \mathcal{P}^1)\right]}{\left[\sum_{i=1}^n S_i^1 \lambda_C(Z_i^1 \mid \mathcal{P}^2)\right]/\left[\sum_{i=1}^n S_i^1 \lambda(Z_i^1 \mid \mathcal{P}^2)\right]} \times \frac{\left[\sum_{i=1}^n S_i^2 \lambda_C(Z_i^2 \mid \mathcal{P}^2)\right]/\left[\sum_{i=1}^n S_i^2 \lambda(Z_i^2 \mid \mathcal{P}^2)\right]}{\left[\sum_{i=1}^n S_i^1 \lambda_C(Z_i^1 \mid \mathcal{P}^1)\right]/\left[\sum_{i=1}^n S_i^1 \lambda(Z_i^1 \mid \mathcal{P}^1)\right]}\right)^{-1/2}}_{\overline{\mathcal{S}}_2}.\end{aligned} \tag{2.12}$$

Here, $\overline{\mathcal{T}}_2$ provides an aggregate measure of technological change from time 1 to time 2, irrespective of whether CRS or VRS holds for the technology. Meanwhile, $\overline{\mathcal{S}}_2$ offers an alternative measure of the aggregated change in scale efficiency; see Ray and Desli (1997) and Färe et al. (1997) for further discussion.

Gilbert and Wilson (1998) propose another decomposition of the individual MPI; by adapting the ideas from Zelenyuk (2006), its aggregate analogue is given by

$$\begin{aligned}\overline{\mathcal{M}} := \overline{\mathcal{E}}_2 \times \overline{\mathcal{T}}_2 \times \overline{\mathcal{S}}_1 \times \\ \underbrace{\left(\frac{\left[\sum_{i=1}^n S_i^1 \lambda_C(Z_i^1 \mid \mathcal{P}^1)\right]/\left[\sum_{i=1}^n S_i^1 \lambda(Z_i^1 \mid \mathcal{P}^1)\right]}{\left[\sum_{i=1}^n S_i^1 \lambda_C(Z_i^1 \mid \mathcal{P}^2)\right]/\left[\sum_{i=1}^n S_i^1 \lambda(Z_i^1 \mid \mathcal{P}^2)\right]} \times \frac{\left[\sum_{i=1}^n S_i^2 \lambda_C(Z_i^2 \mid \mathcal{P}^1)\right]/\left[\sum_{i=1}^n S_i^2 \lambda(Z_i^2 \mid \mathcal{P}^1)\right]}{\left[\sum_{i=1}^n S_i^2 \lambda_C(Z_i^2 \mid \mathcal{P}^2)\right]/\left[\sum_{i=1}^n S_i^2 \lambda(Z_i^2 \mid \mathcal{P}^2)\right]}\right)^{-1/2}}_{\overline{\mathcal{S}}_3},\end{aligned} \tag{2.13}$$

where $\overline{\mathcal{S}}_3$ represents the aggregate counterpart of the "ΔScaleTech" measure introduced by Gilbert and Wilson (1998) (see also Simar and Wilson, 1998a and Wheelock and Wilson, 1999), and $\overline{\mathcal{S}}_2 = \overline{\mathcal{S}}_1 \times \overline{\mathcal{S}}_3$. For further details on this measure, see the discussion in the aforementioned references as well as the more recent study by Simar and Wilson (2019).

However, note that $\overline{\mathcal{T}}_2$, $\overline{\mathcal{S}}_2$, and $\overline{\mathcal{S}}_3$ may suffer from infeasibility issues, since the intertime output-oriented technical efficiency measures $\lambda(Z_i^1 \mid \mathcal{P}^2)$ and $\lambda(Z_i^2 \mid \mathcal{P}^1)$ may fail to admit a solution. In our empirical illustration, we encountered infeasibility problems and therefore do not report the estimates of $\overline{\mathcal{T}}_2$, $\overline{\mathcal{S}}_2$, and $\overline{\mathcal{S}}_3$. One potential solution is to define an aggregate MPI using hyperbolic efficiency measures grounded in economic theory (e.g., as in Zelenyuk (2006)). and to develop the corresponding analogous decompositions for this alternative metric. We leave this for future study.

2.4 Reformulation of Aggregate Sources

To facilitate derivations, analogous to Pham et al. (2024), let us introduce auxiliary variables $U_{1,i}, ..., U_{10,i}$ that correspond to weighted efficiency terms appearing in the decomposition:

$$\begin{aligned}
&U_{1,i} = \lambda_C(Z_i^2 \mid \mathcal{P}^1)w_y^2 Y_i^2,\ U_{2,i} = \lambda_C(Z_i^2 \mid \mathcal{P}^2)w_y^2 Y_i^2,\\
&U_{3,i} = \lambda_C(Z_i^1 \mid \mathcal{P}^1)w_y^1 Y_i^1,\ U_{4,i} = \lambda_C(Z_i^1 \mid \mathcal{P}^2)w_y^1 Y_i^1,\\
&U_{5,i} = \lambda(Z_i^2 \mid \mathcal{P}^1)w_y^2 Y_i^2,\ U_{6,i} = \lambda(Z_i^2 \mid \mathcal{P}^2)w_y^2 Y_i^2,\\
&U_{7,i} = \lambda(Z_i^1 \mid \mathcal{P}^1)w_y^1 Y_i^1,\ U_{8,i} = \lambda(Z_i^1 \mid \mathcal{P}^2)w_y^1 Y_i^1,\\
&U_{9,i} = w_y^1 Y_i^1,\ U_{10,i} = w_y^2 Y_i^2,
\end{aligned} \tag{2.14}$$

and denote

$$\mu_s = E(U_{s,i}), \tag{2.15}$$

and

$$\overline{U}_{s,n} = \frac{1}{n}\sum_{i=1}^{n} U_{s,i},\ s = 1, 2, \ldots, 10. \tag{2.16}$$

Therefore, the log version of the components for the aggregate MPI can be denoted as

$$\mathcal{E}_{1,n} := \log \overline{\mathcal{E}}_{1,n} = -(\log \overline{U}_{2,n} - \log \overline{U}_{3,n} + \log \overline{U}_{9,n} - \log \overline{U}_{10,n}), \tag{2.17}$$

$$\mathcal{T}_{1,n} := \log \overline{\mathcal{T}}_{1,n} = -\frac{1}{2}(\log \overline{U}_{1,n} - \log \overline{U}_{2,n} + \log \overline{U}_{3,n} - \log \overline{U}_{4,n}), \tag{2.18}$$

$$\mathcal{E}_{2,n} := \log \overline{\mathcal{E}}_{2,n} = -(\log \overline{U}_{6,n} - \log \overline{U}_{7,n} + \log \overline{U}_{9,n} - \log \overline{U}_{10,n}), \tag{2.19}$$

$$\mathcal{S}_{1,n} := \log \overline{\mathcal{S}}_{1,n} = -(\log \overline{U}_{2,n} - \log \overline{U}_{6,n} + \log \overline{U}_{7,n} - \log \overline{U}_{3,n}), \tag{2.20}$$

$$\mathcal{T}_{2,n} := \log \overline{\mathcal{T}}_{2,n} = -\frac{1}{2}(\log \overline{U}_{5,n} - \log \overline{U}_{6,n} + \log \overline{U}_{7,n} - \log \overline{U}_{8,n}), \tag{2.21}$$

$$\begin{aligned}
\mathcal{S}_{2,n} := \log \overline{\mathcal{S}}_{2,n} = -\frac{1}{2}\big(&\log \overline{U}_{1,n} - \log \overline{U}_{4,n} + \log \overline{U}_{8,n} - \log \overline{U}_{5,n}\\
&+ \log \overline{U}_{2,n} - \log \overline{U}_{3,n} + \log \overline{U}_{7,n} - \log \overline{U}_{6,n}\big),
\end{aligned} \tag{2.22}$$

and

$$\begin{aligned}
\mathcal{S}_{3,n} := \log \overline{\mathcal{S}}_{3,n} = -\frac{1}{2}\big(&\log \overline{U}_{3,n} - \log \overline{U}_{4,n} + \log \overline{U}_{8,n} - \log \overline{U}_{7,n}\\
&+ \log \overline{U}_{1,n} - \log \overline{U}_{2,n} + \log \overline{U}_{6,n} - \log \overline{U}_{5,n}\big),
\end{aligned} \tag{2.23}$$

which are point estimates of

$$\mathcal{E}_1 = -(\log \mu_2 - \log \mu_3 + \log \mu_9 - \log \mu_{10}), \tag{2.24}$$

$$\mathcal{T}_1 = -\frac{1}{2}(\log \mu_1 - \log \mu_2 + \log \mu_3 - \log \mu_4), \tag{2.25}$$

$$\mathcal{E}_2 = -(\log \mu_6 - \log \mu_7 + \log \mu_9 - \log \mu_{10}), \tag{2.26}$$

$$\mathcal{S}_1 = -(\log \mu_2 - \log \mu_6 + \log \mu_7 - \log \mu_3), \tag{2.27}$$

$$\mathcal{T}_2 = -\frac{1}{2}(\log \mu_5 - \log \mu_6 + \log \mu_7 - \log \mu_8), \tag{2.28}$$

$$\begin{aligned} \mathcal{S}_2 = -\frac{1}{2}\big(&\log \mu_1 - \log \mu_4 + \log \mu_8 - \log \mu_5 \\ &+ \log \mu_2 - \log \mu_3 + \log \mu_7 - \log \mu_6\big), \end{aligned} \tag{2.29}$$

and

$$\begin{aligned} \mathcal{S}_3 = -\frac{1}{2}\big(&\log \mu_3 - \log \mu_4 + \log \mu_8 - \log \mu_7 \\ &+ \log \mu_1 - \log \mu_2 + \log \mu_6 - \log \mu_5\big), \end{aligned} \tag{2.30}$$

respectively.

3 DEA-Based Estimation of Individual and Aggregate Sources

The objects described in the previous section are theoretical constructs derived from the economic theory of production. They represent true but unobservable quantities and therefore must be estimated empirically. Several methods can be used for estimation, including data envelopment analysis (DEA), stochastic frontier analysis (SFA), and various hybrid methods that combine features of both. In this study, we concentrate on the DEA framework, although similar developments could be carried out using alternative estimation techniques.

Given a random sample $\mathcal{X}_n$ defined in Subsection 2.2, the values of $\lambda(x, y \mid \mathcal{P}^t)$ can be obtained through variable returns to scale DEA (VRS-DEA) using the following formulation:

$$\widehat{\lambda}(x, y \mid \mathcal{X}_n^t) = \max_{\lambda, \pi_1, \ldots, \pi_n} \Big\{ \lambda \mid \lambda y \leq \sum_{i=1}^{n} \pi_i Y_i^t,\ x \geq \sum_{i=1}^{n} \pi_i X_i^t,\ \sum_{i=1}^{n} \pi_i = 1,\ \lambda \geq 0,\ \forall\ \pi_i \geq 0 \Big\}. \tag{3.1}$$

The values of $\lambda_C(x, y \mid \mathcal{P}^t)$ can be obtained through

$$\widehat{\lambda}_C(x, y \mid \mathcal{X}_n^t) = \max_{\lambda, \pi_1, \ldots, \pi_n} \Big\{ \lambda \mid \lambda y \leq \sum_{i=1}^{n} \pi_i Y_i^t,\ x \geq \sum_{i=1}^{n} \pi_i X_i^t,\ \lambda \geq 0, \forall\ \pi_i \geq 0 \Big\}. \tag{3.2}$$

Plugging the various estimates into the source components defined above in Section 2, we will obtain the corresponding estimates. The $\mathcal{E}_1$, $\mathcal{T}_1$, $\mathcal{E}_2$, $\mathcal{S}_1$, $\mathcal{T}_2$, $\mathcal{S}_2$, and $\mathcal{S}_3$ can then be estimated by

$$\widehat{\mathcal{E}}_{1,n} = -(\log \widehat{\mu}_{2,n} - \log \widehat{\mu}_{3,n} + \log \widehat{\mu}_{9,n} - \log \widehat{\mu}_{10,n}), \tag{3.3}$$

$$\widehat{\mathcal{T}}_{1,n} = -\frac{1}{2}(\log \widehat{\mu}_{1,n} - \log \widehat{\mu}_{2,n} + \log \widehat{\mu}_{3,n} - \log \widehat{\mu}_{4,n}), \tag{3.4}$$

$$\widehat{\mathcal{E}}_{2,n} = -(\log \widehat{\mu}_{6,n} - \log \widehat{\mu}_{7,n} + \log \widehat{\mu}_{9,n} - \log \widehat{\mu}_{10,n}), \tag{3.5}$$

$$\widehat{\mathcal{S}}_{1,n} = -(\log \widehat{\mu}_{2,n} - \log \widehat{\mu}_{6,n} + \log \widehat{\mu}_{7,n} - \log \widehat{\mu}_{3,n}), \tag{3.6}$$

$$\widehat{\mathcal{T}}_{2,n} = -\frac{1}{2}(\log \widehat{\mu}_{5,n} - \log \widehat{\mu}_{6,n} + \log \widehat{\mu}_{7,n} - \log \widehat{\mu}_{8,n}), \tag{3.7}$$

$$\begin{aligned} \widehat{\mathcal{S}}_{2,n} = -\frac{1}{2}\big(& \log \widehat{\mu}_{1,n} - \log \widehat{\mu}_{4,n} + \log \widehat{\mu}_{8,n} - \log \widehat{\mu}_{5,n} \\ & + \log \widehat{\mu}_{2,n} - \log \widehat{\mu}_{3,n} + \log \widehat{\mu}_{7,n} - \log \widehat{\mu}_{6,n}\big), \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} \widehat{\mathcal{S}}_{3,n} = -\frac{1}{2}\big(& \log \widehat{\mu}_{3,n} - \log \widehat{\mu}_{4,n} + \log \widehat{\mu}_{8,n} - \log \widehat{\mu}_{7,n} \\ & + \log \widehat{\mu}_{1,n} - \log \widehat{\mu}_{2,n} + \log \widehat{\mu}_{6,n} - \log \widehat{\mu}_{5,n}\big), \end{aligned} \tag{3.9}$$

respectively, where

$$\widehat{\mu}_{s,n} = \frac{1}{n} \sum_{i=1}^{n} \widehat{U}_{s,i},\ s = 1, 2, \ldots, 8, \tag{3.10}$$

and

$$\widehat{\mu}_{r,n} = \overline{U}_{r,n} = \frac{1}{n}\sum_{i=1}^{n} U_{r,i},\ r = 9, 10, \tag{3.11}$$

and where

$$\begin{aligned}
&\widehat{U}_{1,i} = \widehat{\lambda}_C(Z_i^2 \mid \mathcal{X}_n^1) w_y^2 Y_i^2,\ \widehat{U}_{2,i} = \widehat{\lambda}_C(Z_i^2 \mid \mathcal{X}_n^2) w_y^2 Y_i^2,\\
&\widehat{U}_{3,i} = \widehat{\lambda}_C(Z_i^1 \mid \mathcal{X}_n^1) w_y^1 Y_i^1,\ \widehat{U}_{4,i} = \widehat{\lambda}_C(Z_i^1 \mid \mathcal{X}_n^2) w_y^1 Y_i^1,\\
&\widehat{U}_{5,i} = \widehat{\lambda}(Z_i^2 \mid \mathcal{X}_n^1) w_y^2 Y_i^2,\ \widehat{U}_{6,i} = \widehat{\lambda}(Z_i^2 \mid \mathcal{X}_n^2) w_y^2 Y_i^2,\\
&\widehat{U}_{7,i} = \widehat{\lambda}(Z_i^1 \mid \mathcal{X}_n^1) w_y^1 Y_i^1,\ \widehat{U}_{8,i} = \widehat{\lambda}(Z_i^1 \mid \mathcal{X}_n^2) w_y^1 Y_i^1,\\
&U_{9,i} = w_y^1 Y_i^1,\ U_{10,i} = w_y^2 Y_i^2.
\end{aligned} \tag{3.12}$$

4 Asymptotic Theory

In this section, we provide the asymptotic theory for the estimators described above by building on results from Korostelev et al. (1995), Kneip et al. (1998), Kneip et al. (2015, 2016, 2021), Simar and Zelenyuk (2018), Simar and Wilson (2019), Pham et al. (2024), etc.[8] We will focus on the case where the production technology satisfies variable returns to scale.

4.1 The Main Theoretical Results

Before stating the main result, recall that for the production technology that is assumed to be convex and satisfy variable returns to scale, the recent studies (e.g., Simar and Wilson, 2019; Kneip et al., 2021; Pham et al., 2024) have shown that $\widehat{\lambda}(x, y \mid \mathcal{X}_n^t)$ and $\widehat{\lambda}_C(x, y \mid \mathcal{X}_n^t)$ are consistent with the rate $\kappa = 2/(p + q + 1)$.[9]

For the sake of brevity, we state all the CLT results for these estimators in one theorem below, where we denote $\widehat{\Gamma}_n$ to represent either one of the estimators from $\widehat{\mathcal{E}}_{1,n}$, $\widehat{\mathcal{E}}_{2,n}$, $\widehat{\mathcal{T}}_{1,n}$, $\widehat{\mathcal{T}}_{2,n}$, $\widehat{\mathcal{S}}_{1,n}$, $\widehat{\mathcal{S}}_{2,n}$, and $\widehat{\mathcal{S}}_{3,n}$.

[8] The proofs are omitted for brevity. Interested readers can refer to the procedures outlined in Pham et al. (2024) for analogous details.

[9] It is worth noting that when the technology is convex and exhibits CRS, the decomposition of the aggregate MPI contains only two components, namely $\mathcal{E}_{1,n}$ and $\mathcal{T}_{1,n}$. Some theoretical results presented below continue to hold for these two terms, but with the convergence rate $\kappa = 2/(p + q)$.

Theorem 1. *Under appropriate assumptions, for $\kappa \geq 2/5$ (i.e., $p+q \leq 4$ for VRS case),*

$$\sqrt{n}\left(\widehat{\Gamma}_n - Bias_{\Gamma,n,\kappa} - \Gamma - R_{\Gamma,n,\kappa}\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\Gamma}), \tag{4.1}$$

and if $\kappa < 1/2$ (i.e., $p+q \geq 4$ for VRS case), we have

$$\sqrt{n_\kappa}\left(\widehat{\Gamma}_{n_\kappa} - Bias_{\Gamma,n,\kappa} - \Gamma - R_{\Gamma,n,\kappa}\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\Gamma}), \tag{4.2}$$

where $R_{\Gamma,n,\kappa} = O(n^{-\frac{3}{2}\kappa}(\log n)^{\frac{3}{2}\kappa}) = o(n^{-\kappa})$, σ^2_Γ is the true variance, $Bias_{\Gamma,n,\kappa}$ is the bias of $\widehat{\Gamma}_n$ in estimating Γ, and $\widehat{\Gamma}_{n_\kappa}$ is a suitable random subsample analogue of $\widehat{\Gamma}_n$, where the subsample size is given by $n_\kappa = \lfloor n^{2\kappa} \rfloor < n$.

It is worth noting here that, as in the previously established results, the rate of convergence in this CLT depends on dimensionality of the production model, i.e., $(p+q)$, reflecting the curse of dimensionality of the DEA estimator.

More specifically, a suitable random subsample analogue of $\widehat{\Gamma}_n$, which we denoted by $\widehat{\Gamma}_{n_\kappa}$, can be obtained for each of the estimators of interest as follows:

$$\widehat{\mathcal{E}}_{1,n_\kappa} = -(\log \widehat{\mu}_{2,n_\kappa} - \log \widehat{\mu}_{3,n_\kappa} + \log \widehat{\mu}_{9,n_\kappa} - \log \widehat{\mu}_{10,n_\kappa}), \tag{4.3}$$

$$\widehat{\mathcal{T}}_{1,n_\kappa} = -\frac{1}{2}(\log \widehat{\mu}_{1,n_\kappa} - \log \widehat{\mu}_{2,n_\kappa} + \log \widehat{\mu}_{3,n_\kappa} - \log \widehat{\mu}_{4,n_\kappa}), \tag{4.4}$$

$$\widehat{\mathcal{E}}_{2,n_\kappa} = -(\log \widehat{\mu}_{6,n_\kappa} - \log \widehat{\mu}_{7,n_\kappa} + \log \widehat{\mu}_{9,n_\kappa} - \log \widehat{\mu}_{10,n_\kappa}), \tag{4.5}$$

$$\widehat{\mathcal{S}}_{1,n_\kappa} = -(\log \widehat{\mu}_{2,n_\kappa} - \log \widehat{\mu}_{6,n_\kappa} + \log \widehat{\mu}_{7,n_\kappa} - \log \widehat{\mu}_{3,n_\kappa}), \tag{4.6}$$

$$\widehat{\mathcal{T}}_{2,n_\kappa} = -\frac{1}{2}(\log \widehat{\mu}_{5,n_\kappa} - \log \widehat{\mu}_{6,n_\kappa} + \log \widehat{\mu}_{7,n_\kappa} - \log \widehat{\mu}_{8,n_\kappa}), \tag{4.7}$$

$$\begin{aligned}\widehat{\mathcal{S}}_{2,n_\kappa} = -\frac{1}{2}(&\log \widehat{\mu}_{1,n_\kappa} - \log \widehat{\mu}_{4,n_\kappa} + \log \widehat{\mu}_{8,n_\kappa} - \log \widehat{\mu}_{5,n_\kappa} \\ &+ \log \widehat{\mu}_{2,n_\kappa} - \log \widehat{\mu}_{3,n_\kappa} + \log \widehat{\mu}_{7,n_\kappa} - \log \widehat{\mu}_{6,n_\kappa}),\end{aligned} \tag{4.8}$$

and

$$\begin{aligned}\widehat{\mathcal{S}}_{3,n_\kappa} = -\frac{1}{2}(&\log \widehat{\mu}_{3,n_\kappa} - \log \widehat{\mu}_{4,n_\kappa} + \log \widehat{\mu}_{8,n_\kappa} - \log \widehat{\mu}_{7,n_\kappa} \\ &+ \log \widehat{\mu}_{1,n_\kappa} - \log \widehat{\mu}_{2,n_\kappa} + \log \widehat{\mu}_{6,n_\kappa} - \log \widehat{\mu}_{5,n_\kappa}),\end{aligned} \tag{4.9}$$

where

$$\widehat{\mu}_{s,n_\kappa} = \frac{1}{n_\kappa} \sum_{\{i|(Z_i^1,Z_i^2)\in\mathcal{X}_{n_\kappa}\}} \widehat{U}_{s,i}, \;\; s = 1, 2, \ldots, 8, \tag{4.10}$$

and

$$\widehat{\mu}_{r,n_\kappa} = \frac{1}{n_\kappa} \sum_{\{i|(Z_i^1,Z_i^2)\in\mathcal{X}_{n_\kappa}\}} U_{r,i}, \;\; r = 9, 10, \tag{4.11}$$

and where $\mathcal{X}_{n_\kappa}$ is a random subsample from $\mathcal{X}_n$.[10]

The estimates of the variance and bias term are discussed later in the following subsections. Suppose we already obtained a consistent estimate of σ^2_Γ, denoted as $\widehat{\sigma}^2_\Gamma$, and a consistent estimator for $Bias_{\Gamma,n,\kappa}$, which we denote as $\widehat{B}_{\Gamma,n,\kappa,B}$, then the symmetric $100(1-\alpha)\%$ confidence intervals for Γ can be approximated by[11]

$$\left[\widehat{\Gamma}_n - \widehat{B}_{\Gamma,n,\kappa,B} \;\pm\; \Phi^{-1}_{1-\alpha/2}\, \widehat{\sigma}_\Gamma/\sqrt{n}\right], \tag{4.12}$$

if $\kappa \geq 1/2$; but if $\kappa < 1/2$, then

$$\left[\widehat{\Gamma}_{n_\kappa} - \widehat{B}_{\Gamma,n,\kappa,B} \;\pm\; \Phi^{-1}_{1-\alpha/2}\, \widehat{\sigma}_\Gamma/\sqrt{n_\kappa}\right]. \tag{4.13}$$

4.2 Estimating the Bias

It is well known that nonparametric frontier estimators, such as DEA, suffer from finite-sample bias due to the fact that the estimated production frontier is constructed from observed best-performing units (Simar and Wilson, 1998b). In finite samples, this leads to an overly optimistic estimate of the frontier. As a result, DEA estimates of Farrell efficiency are biased—downward for output-oriented measures (and upward for input-oriented measures). When these efficiency estimates enter the formulas for the MPI and its components, their bias propagates into the resulting productivity measures and their decompositions. This explains the presence of the bias term $Bias_{\Gamma,n,\kappa}$ in the asymptotic results above, which must be accounted for in order for the central limit theorem to hold and for statistical inference—such

[10] We emphasize that $\widehat{U}_{s,i}$ for $s \in \{1, 2, \ldots, 8\}$ (as specified in (3.12)) is calculated for each index i in the subsample of size n_κ, with computations referencing the full dataset $\mathcal{X}_n$ (not the subsampled dataset $\mathcal{X}_{n_\kappa}$). Subsequent averaging of these $\widehat{U}_{s,i}$ values across all i in the subsample produces $\widehat{\mu}_{s,n_\kappa}$.

[11] Analogous to Pham et al. (2024), if $\kappa = 2/5$ then both (4.12) and (4.13) are applicable, but (4.13) is expected to perform better due to a smaller remainder term, as $\sqrt{n_\kappa} R_{\Gamma,n,\kappa}$ converges to zero more quickly than $\sqrt{n} R_{\Gamma,n,\kappa}$.

as the construction of confidence intervals—to be accurate.

In this subsection we briefly describe a consistent estimator for $Bias_{\Gamma,n,\kappa}$, which we denote as $\widehat{B}_{\Gamma,n,\kappa,B}$. To do so, we adapt ideas from earlier studies (especially Simar and Wilson, 2019 and Pham et al., 2024) for employing the generalized jackknife method to consistently estimate the bias inherent in $\widehat{\Gamma}_n$, as described below.

For each $b = 1, 2, \ldots, B$, where B is much smaller than the binomial coefficient $\binom{n}{\lfloor n/2 \rfloor}$, we randomly partition the sample $\mathcal{X}_n$ into two disjoint subsamples $\mathcal{X}_{n/2,1,b}$ and $\mathcal{X}_{n/2,2,b}$ of equal size. We assume n is even for simplicity. These two subsamples satisfy $\mathcal{X}_{n/2,1,b} \cap \mathcal{X}_{n/2,2,b} = \emptyset$ and $\mathcal{X}_{n/2,1,b} \cup \mathcal{X}_{n/2,2,b} = \mathcal{X}_n$. Further, for each $j = 1, 2$, let $\mathcal{X}^1_{n/2,j,b}$ and $\mathcal{X}^2_{n/2,j,b}$ be the observations in $\mathcal{X}_{n/2,j,b}$, split by periods 1 and 2, respectively. For each $j = 1, 2$ and $b = 1, 2, \ldots, B$, compute

$$\begin{aligned}
\widehat{\mu}_{1,j,b} &= \frac{2}{n} \sum_{\{i \mid (Z_i^1, Z_i^2) \in \mathcal{X}_{n/2,j,b}\}} \widehat{\lambda}_C(Z_i^2 \mid \mathcal{X}^1_{n/2,j,b}) w_y^2 Y_i^2, \\
\widehat{\mu}_{2,j,b} &= \frac{2}{n} \sum_{\{i \mid (Z_i^1, Z_i^2) \in \mathcal{X}_{n/2,j,b}\}} \widehat{\lambda}_C(Z_i^2 \mid \mathcal{X}^2_{n/2,j,b}) w_y^2 Y_i^2, \\
\widehat{\mu}_{3,j,b} &= \frac{2}{n} \sum_{\{i \mid (Z_i^1, Z_i^2) \in \mathcal{X}_{n/2,j,b}\}} \widehat{\lambda}_C(Z_i^1 \mid \mathcal{X}^1_{n/2,j,b}) w_y^1 Y_i^1, \\
\widehat{\mu}_{4,j,b} &= \frac{2}{n} \sum_{\{i \mid (Z_i^1, Z_i^2) \in \mathcal{X}_{n/2,j,b}\}} \widehat{\lambda}_C(Z_i^1 \mid \mathcal{X}^2_{n/2,j,b}) w_y^1 Y_i^1, \\
\widehat{\mu}_{5,j,b} &= \frac{2}{n} \sum_{\{i \mid (Z_i^1, Z_i^2) \in \mathcal{X}_{n/2,j,b}\}} \widehat{\lambda}(Z_i^2 \mid \mathcal{X}^1_{n/2,j,b}) w_y^2 Y_i^2, \\
\widehat{\mu}_{6,j,b} &= \frac{2}{n} \sum_{\{i \mid (Z_i^1, Z_i^2) \in \mathcal{X}_{n/2,j,b}\}} \widehat{\lambda}(Z_i^2 \mid \mathcal{X}^2_{n/2,j,b}) w_y^2 Y_i^2, \\
\widehat{\mu}_{7,j,b} &= \frac{2}{n} \sum_{\{i \mid (Z_i^1, Z_i^2) \in \mathcal{X}_{n/2,j,b}\}} \widehat{\lambda}(Z_i^1 \mid \mathcal{X}^1_{n/2,j,b}) w_y^1 Y_i^1, \\
\widehat{\mu}_{8,j,b} &= \frac{2}{n} \sum_{\{i \mid (Z_i^1, Z_i^2) \in \mathcal{X}_{n/2,j,b}\}} \widehat{\lambda}(Z_i^1 \mid \mathcal{X}^2_{n/2,j,b}) w_y^1 Y_i^1,
\end{aligned} \tag{4.14}$$

and compute the subsample version $\widehat{\Gamma}_{n/2,j,b}$. For example, if $\widehat{\Gamma}_n = \widehat{\mathcal{E}}_{1,n}$, then

$$\widehat{\Gamma}_{n/2,j,b} = \widehat{\mathcal{E}}_{1,n/2,j,b} = -(\log \widehat{\mu}_{2,j,b} - \log \widehat{\mu}_{3,j,b} + \log \widehat{\mu}_{9,n} - \log \widehat{\mu}_{10,n}). \tag{4.15}$$

In essence, $\widehat{\mathcal{E}}_{1,n/2,j,b}$ serves as an analogue of $\widehat{\mathcal{E}}_{1,n}$, where the components involving efficiency scores are evaluated over the subsample $\mathcal{X}_{n/2,j,b}$, while the remaining components (i.e., $\widehat{\mu}_{9,n}$ and $\widehat{\mu}_{10,n}$) are evaluated using the full sample $\mathcal{X}_n$. Then compute

$$\widehat{\mathcal{E}}^*_{1,n/2,b} = \frac{1}{2}(\widehat{\mathcal{E}}_{1,n/2,1,b} + \widehat{\mathcal{E}}_{1,n/2,2,b}). \tag{4.16}$$

An estimate of the bias term associated with $\widehat{\mathcal{E}}_{1,n}$ is given by

$$\widehat{B}_{\mathcal{E}_{1,n,\kappa,B}} = \frac{1}{B}\sum_{b=1}^{B}(2^\kappa - 1)^{-1}(\widehat{\mathcal{E}}^*_{1,n/2,b} - \widehat{\mathcal{E}}_{1,n}). \tag{4.17}$$

If $\widehat{\Gamma}_n = \widehat{\mathcal{T}}_{1,n}$, then

$$\widehat{\Gamma}_{n/2,j,b} = \widehat{\mathcal{T}}_{1,n/2,j,b} = -\frac{1}{2}(\log\widehat{\mu}_{1,j,b} - \log\widehat{\mu}_{2,j,b} + \log\widehat{\mu}_{3,j,b} - \log\widehat{\mu}_{4,j,b}). \tag{4.18}$$

Then compute

$$\widehat{\mathcal{T}}^*_{1,n/2,b} = \frac{1}{2}(\widehat{\mathcal{T}}_{1,n/2,1,b} + \widehat{\mathcal{T}}_{1,n/2,2,b}). \tag{4.19}$$

An estimate of the bias term associated with $\widehat{T}_{1,n}$ is given by

$$\widehat{B}_{\mathcal{T}_{1,n,\kappa,B}} = \frac{1}{B}\sum_{b=1}^{B}(2^\kappa - 1)^{-1}(\widehat{\mathcal{T}}^*_{1,n/2,b} - \widehat{\mathcal{T}}_{1,n}). \tag{4.20}$$

If $\widehat{\Gamma}_n = \widehat{\mathcal{E}}_{2,n}$, then

$$\widehat{\Gamma}_{n/2,j,b} = \widehat{\mathcal{E}}_{2,n/2,j,b} = -(\log\widehat{\mu}_{6,j,b} - \log\widehat{\mu}_{7,j,b} + \log\widehat{\mu}_{9,n} - \log\widehat{\mu}_{10,n}). \tag{4.21}$$

Then compute

$$\widehat{\mathcal{E}}^*_{2,n/2,b} = \frac{1}{2}(\widehat{\mathcal{E}}_{2,n/2,1,b} + \widehat{\mathcal{E}}_{2,n/2,2,b}). \tag{4.22}$$

An estimate of the bias term associated with $\widehat{\mathcal{E}}_{2,n}$ is given by

$$\widehat{B}_{\mathcal{E}_{2,n,\kappa,B}} = \frac{1}{B}\sum_{b=1}^{B}(2^{\kappa}-1)^{-1}(\widehat{\mathcal{E}}^{*}_{2,n/2,b} - \widehat{\mathcal{E}}_{2,n}). \tag{4.23}$$

If $\widehat{\Gamma}_n = \widehat{\mathcal{S}}_{1,n}$, then

$$\widehat{\Gamma}_{n/2,j,b} = \widehat{\mathcal{S}}_{1,n/2,j,b} = -(\log \widehat{\mu}_{2,j,b} - \log \widehat{\mu}_{6,j,b} + \log \widehat{\mu}_{7,j,b} - \log \widehat{\mu}_{3,j,b}). \tag{4.24}$$

Then compute

$$\widehat{\mathcal{S}}^{*}_{1,n/2,b} = \frac{1}{2}(\widehat{\mathcal{S}}_{1,n/2,1,b} + \widehat{\mathcal{S}}_{1,n/2,2,b}). \tag{4.25}$$

An estimate of the bias term associated with $\widehat{\mathcal{S}}_{1,n}$ is given by

$$\widehat{B}_{\mathcal{S}_{1,n,\kappa,B}} = \frac{1}{B}\sum_{b=1}^{B}(2^{\kappa}-1)^{-1}(\widehat{\mathcal{S}}^{*}_{1,n/2,b} - \widehat{\mathcal{S}}_{1,n}). \tag{4.26}$$

If $\widehat{\Gamma}_n = \widehat{\mathcal{T}}_{2,n}$, then

$$\widehat{\Gamma}_{n/2,j,b} = \widehat{\mathcal{T}}_{2,n/2,j,b} = -\frac{1}{2}(\log \widehat{\mu}_{5,j,b} - \log \widehat{\mu}_{6,j,b} + \log \widehat{\mu}_{7,j,b} - \log \widehat{\mu}_{8,j,b}). \tag{4.27}$$

Then compute

$$\widehat{\mathcal{T}}^{*}_{2,n/2,b} = \frac{1}{2}(\widehat{\mathcal{T}}_{2,n/2,1,b} + \widehat{\mathcal{T}}_{2,n/2,2,b}). \tag{4.28}$$

An estimate of the bias term associated with $\widehat{\mathcal{T}}_{2,n}$ is given by

$$\widehat{B}_{\mathcal{T}_{2,n,\kappa,B}} = \frac{1}{B}\sum_{b=1}^{B}(2^{\kappa}-1)^{-1}(\widehat{\mathcal{T}}^{*}_{2,n/2,b} - \widehat{\mathcal{T}}_{2,n}). \tag{4.29}$$

If $\widehat{\Gamma}_n = \widehat{\mathcal{S}}_{2,n}$, then

$$\begin{aligned}\widehat{\Gamma}_{n/2,j,b} = \widehat{\mathcal{S}}_{2,n/2,j,b} = -\frac{1}{2}(&\log \widehat{\mu}_{1,j,b} - \log \widehat{\mu}_{4,j,b} + \log \widehat{\mu}_{8,j,b} - \log \widehat{\mu}_{5,j,b} \\ &+ \log \widehat{\mu}_{2,j,b} - \log \widehat{\mu}_{3,j,b} + \log \widehat{\mu}_{7,j,b} - \log \widehat{\mu}_{6,j,b}).\end{aligned} \tag{4.30}$$

Then compute

$$\widehat{\mathcal{S}}^*_{2,n/2,b} = \frac{1}{2}(\widehat{\mathcal{S}}_{2,n/2,1,b} + \widehat{\mathcal{S}}_{2,n/2,2,b}). \tag{4.31}$$

An estimate of the bias term associated with $\widehat{\mathcal{S}}_{2,n}$ is given by

$$\widehat{B}_{\mathcal{S}_{2},n,\kappa,B} = \frac{1}{B}\sum_{b=1}^{B}(2^{\kappa}-1)^{-1}(\widehat{\mathcal{S}}^*_{2,n/2,b} - \widehat{\mathcal{S}}_{2,n}). \tag{4.32}$$

If $\widehat{\Gamma}_n = \widehat{\mathcal{S}}_{3,n}$, then

$$\begin{aligned}\widehat{\Gamma}_{n/2,j,b} = \widehat{\mathcal{S}}_{3,n/2,j,b} = -\frac{1}{2}\big(&\log\widehat{\mu}_{3,j,b} - \log\widehat{\mu}_{4,j,b} + \log\widehat{\mu}_{8,j,b} - \log\widehat{\mu}_{7,j,b} \\ &+ \log\widehat{\mu}_{1,j,b} - \log\widehat{\mu}_{2,j,b} + \log\widehat{\mu}_{6,j,b} - \log\widehat{\mu}_{5,j,b}\big).\end{aligned} \tag{4.33}$$

Then compute

$$\widehat{\mathcal{S}}^*_{3,n/2,b} = \frac{1}{2}(\widehat{\mathcal{S}}_{3,n/2,1,b} + \widehat{\mathcal{S}}_{3,n/2,2,b}). \tag{4.34}$$

An estimate of the bias term associated with $\widehat{\mathcal{S}}_{3,n}$ is given by

$$\widehat{B}_{\mathcal{S}_{3},n,\kappa,B} = \frac{1}{B}\sum_{b=1}^{B}(2^{\kappa}-1)^{-1}(\widehat{\mathcal{S}}^*_{3,n/2,b} - \widehat{\mathcal{S}}_{3,n}). \tag{4.35}$$

The key asymptotic property of the proposed bias estimator $\widehat{B}_{\Gamma,n,\kappa,B}$ is summarized in the following theorem.[12]

Theorem 2. *Under appropriate assumptions, as $n \to \infty$,*

$$\widehat{B}_{\Gamma,n,\kappa,B} = Bias_{\Gamma,n,\kappa} + O(n^{-\frac{3}{2}\kappa}(\log n)^{\frac{3}{2}\kappa}) + o(n^{-1/2}). \tag{4.36}$$

[12] This result can be obtained by building on the theoretical results in Simar and Wilson (2019), Kneip et al. (2021) and Pham et al. (2024).

4.3 Estimating the Variance

We note that while the true variance σ_Γ^2 is unobserved, it can be consistently estimated using an empirical analog, drawing on the insights presented in Pham et al. (2024), i.e., using

$$\widehat{\sigma}_\Gamma^2 = [\nabla\widehat{\Gamma}_n]'\widehat{\Sigma}_\Gamma[\nabla\widehat{\Gamma}_n], \tag{4.37}$$

where $\nabla\widehat{\Gamma}_n$ represents the column vector corresponding to the gradient of $\widehat{\Gamma}_n$ taken with respect to its elements, and $\widehat{\Sigma}_\Gamma$ is the corresponding covariance matrix. To be more specific, if we consider $\widehat{\Gamma}_n = \widehat{\mathcal{E}}_{1,n}$, then $\nabla\widehat{\Gamma}_n = \nabla\widehat{\mathcal{E}}_{1,n} = [\frac{\partial\widehat{\mathcal{E}}_{1,n}}{\partial\widehat{\mu}_{s,n}}]'_{\{s=2,3,9,10\}}$, where

$$\frac{\partial\widehat{\mathcal{E}}_{1,n}}{\partial\widehat{\mu}_{2,n}} = -\frac{1}{\widehat{\mu}_{2,n}},\ \frac{\partial\widehat{\mathcal{E}}_{1,n}}{\partial\widehat{\mu}_{3,n}} = \frac{1}{\widehat{\mu}_{3,n}},\ \frac{\partial\widehat{\mathcal{E}}_{1,n}}{\partial\widehat{\mu}_{9,n}} = -\frac{1}{\widehat{\mu}_{9,n}},\ \frac{\partial\widehat{\mathcal{E}}_{1,n}}{\partial\widehat{\mu}_{10,n}} = \frac{1}{\widehat{\mu}_{10,n}}, \tag{4.38}$$

and $\widehat{\Sigma}_\Gamma$ is the covariance matrix of $(\widehat{U}_{2,i}, \widehat{U}_{3,i}, U_{9,i}, U_{10,i})$.

If we consider $\widehat{\Gamma}_n = \widehat{\mathcal{T}}_{1,n}$, then $\nabla\widehat{\Gamma}_n = \nabla\widehat{\mathcal{T}}_{1,n} = [\frac{\partial\widehat{\mathcal{T}}_{1,n}}{\partial\widehat{\mu}_{s,n}}]'_{\{s=1,2,3,4\}}$, where

$$\frac{\partial\widehat{\mathcal{T}}_{1,n}}{\partial\widehat{\mu}_{1,n}} = -\frac{1}{2\widehat{\mu}_{1,n}},\ \frac{\partial\widehat{\mathcal{T}}_{1,n}}{\partial\widehat{\mu}_{2,n}} = \frac{1}{2\widehat{\mu}_{2,n}},\ \frac{\partial\widehat{\mathcal{T}}_{1,n}}{\partial\widehat{\mu}_{3,n}} = -\frac{1}{2\widehat{\mu}_{3,n}},\ \frac{\partial\widehat{\mathcal{T}}_{1,n}}{\partial\widehat{\mu}_{4,n}} = \frac{1}{2\widehat{\mu}_{4,n}}, \tag{4.39}$$

and $\widehat{\Sigma}_\Gamma$ is the covariance matrix of $(\widehat{U}_{1,i}, \widehat{U}_{2,i}, \widehat{U}_{3,i}, \widehat{U}_{4,i})$.

If we consider $\widehat{\Gamma}_n = \widehat{\mathcal{E}}_{2,n}$, then $\nabla\widehat{\Gamma}_n = \nabla\widehat{\mathcal{E}}_{2,n} = [\frac{\partial\widehat{\mathcal{E}}_{2,n}}{\partial\widehat{\mu}_{s,n}}]'_{\{s=6,7,9,10\}}$, where

$$\frac{\partial\widehat{\mathcal{E}}_{2,n}}{\partial\widehat{\mu}_{6,n}} = -\frac{1}{\widehat{\mu}_{6,n}},\ \frac{\partial\widehat{\mathcal{E}}_{2,n}}{\partial\widehat{\mu}_{7,n}} = \frac{1}{\widehat{\mu}_{7,n}},\ \frac{\partial\widehat{\mathcal{E}}_{2,n}}{\partial\widehat{\mu}_{9,n}} = -\frac{1}{\widehat{\mu}_{9,n}},\ \frac{\partial\widehat{\mathcal{E}}_{2,n}}{\partial\widehat{\mu}_{10,n}} = \frac{1}{\widehat{\mu}_{10,n}}, \tag{4.40}$$

and $\widehat{\Sigma}_\Gamma$ is the covariance matrix of $(\widehat{U}_{6,i}, \widehat{U}_{7,i}, U_{9,i}, U_{10,i})$.

If we consider $\widehat{\Gamma}_n = \widehat{\mathcal{S}}_{1,n}$, then $\nabla\widehat{\Gamma}_n = \nabla\widehat{\mathcal{S}}_{1,n} = [\frac{\partial\widehat{\mathcal{S}}_{1,n}}{\partial\widehat{\mu}_{s,n}}]'_{\{s=2,6,7,3\}}$, where

$$\frac{\partial\widehat{\mathcal{S}}_{1,n}}{\partial\widehat{\mu}_{2,n}} = -\frac{1}{\widehat{\mu}_{2,n}},\ \frac{\partial\widehat{\mathcal{S}}_{1,n}}{\partial\widehat{\mu}_{6,n}} = \frac{1}{\widehat{\mu}_{6,n}},\ \frac{\partial\widehat{\mathcal{S}}_{1,n}}{\partial\widehat{\mu}_{7,n}} = -\frac{1}{\widehat{\mu}_{7,n}},\ \frac{\partial\widehat{\mathcal{S}}_{1,n}}{\partial\widehat{\mu}_{3,n}} = \frac{1}{\widehat{\mu}_{3,n}}, \tag{4.41}$$

and $\widehat{\Sigma}_\Gamma$ is the covariance matrix of $(\widehat{U}_{2,i}, \widehat{U}_{6,i}, \widehat{U}_{7,i}, \widehat{U}_{3,i})$.

If we consider $\widehat{\Gamma}_n = \widehat{\mathcal{T}}_{2,n}$, then $\nabla\widehat{\Gamma}_n = \nabla\widehat{\mathcal{T}}_{2,n} = [\frac{\partial\widehat{\mathcal{T}}_{2,n}}{\partial\widehat{\mu}_{s,n}}]'_{\{s=5,6,7,8\}}$, where

$$\frac{\partial\widehat{\mathcal{T}}_{2,n}}{\partial\widehat{\mu}_{5,n}} = -\frac{1}{2\widehat{\mu}_{5,n}},\ \frac{\partial\widehat{\mathcal{T}}_{2,n}}{\partial\widehat{\mu}_{6,n}} = \frac{1}{2\widehat{\mu}_{6,n}},\ \frac{\partial\widehat{\mathcal{T}}_{2,n}}{\partial\widehat{\mu}_{7,n}} = -\frac{1}{2\widehat{\mu}_{7,n}},\ \frac{\partial\widehat{\mathcal{T}}_{2,n}}{\partial\widehat{\mu}_{8,n}} = \frac{1}{2\widehat{\mu}_{8,n}}, \tag{4.42}$$

and $\widehat{\Sigma}_\Gamma$ is the covariance matrix of $(\widehat{U}_{5,i}, \widehat{U}_{6,i}, \widehat{U}_{7,i}, \widehat{U}_{8,i})$.

If we consider $\widehat{\Gamma}_n = \widehat{\mathcal{S}}_{2,n}$, then $\nabla\widehat{\Gamma}_n = \nabla\widehat{\mathcal{S}}_{2,n} = [\frac{\partial \widehat{\mathcal{S}}_{2,n}}{\partial \widehat{\mu}_{s,n}}]'_{\{s=1,4,8,5,2,3,7,6\}}$, where

$$\begin{aligned}
&\frac{\partial \widehat{\mathcal{S}}_{2,n}}{\partial \widehat{\mu}_{1,n}} = -\frac{1}{2\widehat{\mu}_{1,n}}, \ \frac{\partial \widehat{\mathcal{S}}_{2,n}}{\partial \widehat{\mu}_{4,n}} = \frac{1}{2\widehat{\mu}_{4,n}}, \ \frac{\partial \widehat{\mathcal{S}}_{2,n}}{\partial \widehat{\mu}_{8,n}} = -\frac{1}{2\widehat{\mu}_{8,n}}, \ \frac{\partial \widehat{\mathcal{S}}_{2,n}}{\partial \widehat{\mu}_{5,n}} = \frac{1}{2\widehat{\mu}_{5,n}}, \\
&\frac{\partial \widehat{\mathcal{S}}_{2,n}}{\partial \widehat{\mu}_{2,n}} = -\frac{1}{2\widehat{\mu}_{2,n}}, \ \frac{\partial \widehat{\mathcal{S}}_{2,n}}{\partial \widehat{\mu}_{3,n}} = \frac{1}{2\widehat{\mu}_{3,n}}, \ \frac{\partial \widehat{\mathcal{S}}_{2,n}}{\partial \widehat{\mu}_{7,n}} = -\frac{1}{2\widehat{\mu}_{7,n}}, \ \frac{\partial \widehat{\mathcal{S}}_{2,n}}{\partial \widehat{\mu}_{6,n}} = \frac{1}{2\widehat{\mu}_{6,n}},
\end{aligned} \tag{4.43}$$

and $\widehat{\Sigma}_\Gamma$ is the covariance matrix of $(\widehat{U}_{1,i}, \widehat{U}_{4,i}, \widehat{U}_{8,i}, \widehat{U}_{5,i}, \widehat{U}_{2,i}, \widehat{U}_{3,i}, \widehat{U}_{7,i}, \widehat{U}_{6,i})$.

If we consider $\widehat{\Gamma}_n = \widehat{\mathcal{S}}_{3,n}$, then $\nabla\widehat{\Gamma}_n = \nabla\widehat{\mathcal{S}}_{3,n} = [\frac{\partial \widehat{\mathcal{S}}_{3,n}}{\partial \widehat{\mu}_{s,n}}]'_{\{s=3,4,8,7,1,2,6,5\}}$, where

$$\begin{aligned}
&\frac{\partial \widehat{\mathcal{S}}_{3,n}}{\partial \widehat{\mu}_{3,n}} = -\frac{1}{2\widehat{\mu}_{3,n}}, \ \frac{\partial \widehat{\mathcal{S}}_{3,n}}{\partial \widehat{\mu}_{4,n}} = \frac{1}{2\widehat{\mu}_{4,n}}, \ \frac{\partial \widehat{\mathcal{S}}_{3,n}}{\partial \widehat{\mu}_{8,n}} = -\frac{1}{2\widehat{\mu}_{8,n}}, \ \frac{\partial \widehat{\mathcal{S}}_{3,n}}{\partial \widehat{\mu}_{7,n}} = \frac{1}{2\widehat{\mu}_{7,n}}, \\
&\frac{\partial \widehat{\mathcal{S}}_{3,n}}{\partial \widehat{\mu}_{1,n}} = -\frac{1}{2\widehat{\mu}_{1,n}}, \ \frac{\partial \widehat{\mathcal{S}}_{3,n}}{\partial \widehat{\mu}_{2,n}} = \frac{1}{2\widehat{\mu}_{2,n}}, \ \frac{\partial \widehat{\mathcal{S}}_{3,n}}{\partial \widehat{\mu}_{6,n}} = -\frac{1}{2\widehat{\mu}_{6,n}}, \ \frac{\partial \widehat{\mathcal{S}}_{3,n}}{\partial \widehat{\mu}_{5,n}} = \frac{1}{2\widehat{\mu}_{5,n}},
\end{aligned} \tag{4.44}$$

and $\widehat{\Sigma}_\Gamma$ is the covariance matrix of $(\widehat{U}_{3,i}, \widehat{U}_{4,i}, \widehat{U}_{8,i}, \widehat{U}_{7,i}, \widehat{U}_{1,i}, \widehat{U}_{2,i}, \widehat{U}_{6,i}, \widehat{U}_{5,i})$.

5 Some Special Cases

In this section, we examine one important and widely used special case in economics, usually referred to as Hicks-neutral technology change. This type of technology is employed in many economic theory models and is useful to consider separately. Interestingly, in this case we find that the CLT-based inference methods for both weighted and unweighted technology changes may not be valid for any dimensions of outputs and inputs and alternative inferential methods are required.

Since we focus on the output orientation, we will consider output Hicks-neutral technology change property here and analogous arguments can be done for the input orientation. Recall that a general definition of this notion (Sickles and Zelenyuk, 2019) states that technology change between some periods s and t is Hicks output neutral if and only if the output correspondence implied by the technology set in period t can be decomposed into the output correspondence implied by the technology set in period s and a scaling function that does not depend on outputs, i.e.,

$$\mathcal{P}_o^t(x) = A_o(x, s, t)\mathcal{P}_o^s(x), \tag{5.1}$$

for some relevant domain of x (usually assumed for all $x \in \mathbb{R}^p_+$), where $A_o(x, s, t)$ is a scaling function that depends only on inputs x and periods s and t, and is independent of outputs, and where $\mathcal{P}^\tau_o(x)$ is the output correspondence in period $\tau \in \{s, t\}$ implied by the technology set $\mathcal{P}^\tau$ in that period, i.e.,

$$\mathcal{P}^\tau_o(x) = \{y \in \mathbb{R}^q_+ : (x, y) \in \mathcal{P}^\tau\}, x \in \mathbb{R}^p_+. \tag{5.2}$$

In its turn, this type of technology implies that the Farrell's output oriented measure of technical efficiency can also be decomposed as (Sickles and Zelenyuk, 2019, p.122):

$$\lambda(x, y \mid \mathcal{P}^t) = A_o(x, s, t)\lambda(x, y \mid \mathcal{P}^s) \tag{5.3}$$

for the relevant domain of x where the Hicks-neutrality is assumed to hold.

Therefore, the technology change component of MPI from period 1 to period 2 for observation i will be

$$\begin{aligned} T_{2,i} &= \left(\frac{\lambda(X^2_i, Y^2_i \mid \mathcal{P}^1)}{\lambda(X^2_i, Y^2_i \mid \mathcal{P}^2)} \times \frac{\lambda(X^1_i, Y^1_i \mid \mathcal{P}^1)}{\lambda(X^1_i, Y^1_i \mid \mathcal{P}^2)} \right)^{-1/2} \\ &= \left(\frac{\lambda(X^2_i, Y^2_i \mid \mathcal{P}^1)}{\lambda(X^2_i, Y^2_i \mid \mathcal{P}^1)A_o(X^2_i, 1, 2)} \times \frac{\lambda(X^1_i, Y^1_i \mid \mathcal{P}^1)}{\lambda(X^1_i, Y^1_i \mid \mathcal{P}^1)A_o(X^1_i, 1, 2)} \right)^{-1/2} \\ &= \left(A_o(X^1_i, 1, 2)A_o(X^2_i, 1, 2) \right)^{1/2}. \end{aligned} \tag{5.4}$$

Thus, for the case when $A_o(x, s, t) = A_o(s, t)$, i.e., the technology change is independent of inputs x (which is the usual case of Hicks-neutrality in economics), the technology change will only depend on the periods compared (s and t) and, importantly, not vary across individual observations i. For a particular choice of s and t, this scaling factor will be some positive constant, which we will denote as A. Hence, the true standard deviation for the technology change will be zero. Note that when $A_o(x, s, t) = A_o(s, t)$, a Hicks-neutral technology change also implies a Hicks-neutral conical technology change and so the true standard deviation for the conical technology change will also be zero.

As a consequence, the standard central limit theorems in Simar and Wilson (2019) for the simple mean case are not applicable. For example, Theorem 4.2 in Simar and Wilson (2019) requires that the true standard deviation is greater than 0, whereas in the Hicks-neutral technology change, the true standard deviation is 0. Analogously, the true standard deviation in the weighted case is also zero, which implies that the theorem developed above cannot be applied.

Another interesting case is the single-output-single-input case. This case is often used in economics textbooks for illustrating geometric intuition and so deserves special attention. To be precise, suppose that the true technologies for both periods are suitable production functions, with $p = 1$ and $q = 1$, satisfying standard regularity conditions of production theory (with any returns to scale assumption and not necessarily Hicks-neutral). Then, the conical technology change measured using any observation is constant. This indicates that the true standard deviation of both the unweighted and weighted conical technology change is zero.[13] Therefore, the standard central limit theorems of both the unweighted and weighted conical technology change are also not applicable here.

In the special cases like described above, whenever technology change between two periods of interest, s and t, is a constant A, we can use alternative methods to conduct statistical inference on A. In particular, to sketch some ideas here, note that in such cases, for a fixed point (x, y), we have

$$A = \frac{\lambda_C(x, y|\mathcal{P}^2)}{\lambda_C(x, y|\mathcal{P}^1)}. \tag{5.5}$$

and its empirical analogue can be obtained as

$$\widehat{A} = \frac{\widehat{\lambda}_C(x, y|\mathcal{X}_n^2)}{\widehat{\lambda}_C(x, y|\mathcal{X}_n^1)}. \tag{5.6}$$

Thus, by a Taylor-series approximation argument, it should be possible to derive a limiting distribution, with finite mean as well as finite and positive variance for $n^\kappa(\widehat{A} - A)$. This would be building on the developments in Park et al. (2010), who established such results for $n^\kappa(\widehat{\lambda}_C(x, y|\mathcal{X}_n^t) - \lambda_C(x, y|\mathcal{P}^t))$. As in the Park et al. (2010), the resulting limiting distribution may be very complicated to apply in practice and so, in principle, bootstrap method can be deployed to construct confidence intervals for A and bias estimation for $\widehat{A}$. For example, one can adapt a consistent bootstrap procedure based on the subsampling approach discussed in Simar and Wilson (2011) or other alternatives. Exploring this topic could be a fruitful

[13] It is worth noting that the technology change here measured using any observation is not constant; therefore, the true standard deviation of both the unweighted and weighted technology change need not be zero.

avenue for further research.

6 Empirical Illustrations

To illustrate our theoretically derived results on the aggregate sources of productivity change, we employ the latest data from the Penn World Table (PWT 11.0). We also compare the empirical results between weighted methods developed in this paper and those obtained using unweighted methods developed by Simar and Wilson (2019).

Consistent with prior literature (e.g., Pham et al., 2024; Zelenyuk and Zhao, 2025), we impose the assumption that the full set of 84 countries/regions employ capital stock (*cn*) and labor (*emp*) as inputs to produce GDP (*rgdpo*) from 1990 to 2023. We conduct analyses for non-overlapping year pairs at 5-year intervals (1990–1995, 1995–2000, ..., 2020–2023) and for the full 1990–2023 period to examine the evolution of productivity change and its sources. Moreover, we set the price of real GDP to be one, so that the weights used are the GDP shares of each country among these 84 countries. Table 2 presents the empirical results obtained using the weighted approach, whereas Table 3 reports the corresponding results based on the unweighted approach. Note that $\widetilde{\Gamma}_n$ denotes the bias-corrected estimate of Γ, whereas $\widehat{\Gamma}_n$ denotes the estimate without bias correction; that is, $\widetilde{\Gamma}_n = \widehat{\Gamma}_n - \widehat{B}_{\Gamma,n,\kappa,B}$.

A first notable result concerns the magnitude of the bias in the uncorrected estimates. Across most rows, the pre-bias-correction estimates differ substantially from the post-bias-correction estimates, underscoring the practical importance of addressing estimator bias in empirical analyses of aggregate productivity change and its sources. Without the bias correction, productivity change and its components may be either over- or under-estimated, and in some cases even the direction of change may be reversed, potentially leading to misleading conclusions.

For example, Table 2 indicates that the weighted productivity for the full sample increased by 22.0% over the period 1990 to 2023, prior to the application of bias correction, compared with a 48.9% increase after correction. Similarly, weighted productivity increased by 0.8% from 2000 to 2005 prior to the bias correction, but after correction it decreased by 3.2% over the same period. Similar discrepancies are also evident for the components of the aggregate MPI, indicating that bias affects not only overall productivity change but also its decomposition.

Table 2: Estimates of Weighted Productivity Change and its Components Across 84 Countries/Regions

Period	$\exp(\widehat{\mathcal{M}}_n)$	$\exp(\widetilde{\mathcal{M}}_n)$	$\exp(\widehat{\mathcal{E}}_{1,n})$	$\exp(\widetilde{\mathcal{E}}_{1,n})$	$\exp(\widehat{\mathcal{E}}_{2,n})$	$\exp(\widetilde{\mathcal{E}}_{2,n})$	$\exp(\widehat{\mathcal{T}}_{1,n})$	$\exp(\widetilde{\mathcal{T}}_{1,n})$	$\exp(\widehat{\mathcal{S}}_{1,n})$	$\exp(\widetilde{\mathcal{S}}_{1,n})$
1990–1995	1.025	1.031	1.188	1.256***	0.975	0.969	0.862	0.821***	1.218	1.296***
1995–2000	1.060	1.039	1.087	1.157**	1.020	1.032**	0.976	0.898	1.065	1.121*
2000–2005	1.008	0.968	0.931	0.943	1.008	1.008	1.083	1.027	0.923	0.935
2005–2010	0.954	0.991	0.980	0.981	1.004	0.971	0.973	1.010	0.976	1.011
2010–2015	1.000	1.020	1.087	1.239***	0.972	0.939***	0.921	0.823***	1.117	1.320***
2015–2020	0.944	0.935***	1.002	0.985	0.953	0.912***	0.942	0.948***	1.051	1.080***
2020–2023	1.021	1.014	0.942	0.895***	0.994	0.980***	1.084	1.132***	0.947	0.914***
1990–2023	1.220	1.489***	1.207	1.509**	0.927	0.811***	1.011	0.987	1.302	1.861***

NOTE: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (deviations from 1). For notational simplicity, we denote $\widetilde{\Gamma}_n = \widehat{\Gamma}_n - \widehat{B}_{\Gamma,n,\kappa,B}$, where $\Gamma \in \{\mathcal{M}, \mathcal{E}_1, \mathcal{E}_2, \mathcal{T}_1, \mathcal{S}_1\}$, i.e., $\widetilde{\Gamma}_n$ denotes the bias-corrected estimate of Γ.

Second, over the period 1990–2023, the bias-corrected estimate of aggregate MPI ($\exp(\widetilde{\mathcal{M}}_n)$) is 1.489 and statistically significant, implying that aggregate productivity across the 84 countries increased by 48.9%. In other words, relative to 1990, the same level of inputs could generate 48.9% more GDP in 2023.

Turning to the decomposition, the bias-corrected estimate of the conical efficiency index ($\exp(\widetilde{\mathcal{E}}_{1,n})$) equals 1.509 and is statistically significant, whereas the bias-corrected estimate of technological change ($\exp(\widetilde{\mathcal{T}}_{1,n})$) is 0.987 and not statistically significant. These findings suggest that the observed aggregate productivity growth is primarily attributable to improvements in conical technical efficiency rather than technological progress.

Furthermore, since the aggregate conical technical efficiency ($\exp(\mathcal{E}_1)$) can be decomposed into the aggregate pure technical efficiency change ($\exp(\mathcal{E}_2)$) and the aggregate scale efficiency change ($\exp(\mathcal{S}_1)$), we examine these components separately. Table 2 reports that, for 1990–2023, the bias-corrected estimate of the aggregate pure technical efficiency change ($\exp(\widetilde{\mathcal{E}}_{2,n})$) is 0.811 and statistically significant, while the bias-corrected estimate of the aggregate scale efficiency change ($\exp(\widetilde{\mathcal{S}}_{1,n})$) is 1.861 and statistically significant. This decomposition indicates that the increase in the aggregate conical technical efficiency is mainly driven by improvements in the aggregate scale efficiency, whereas aggregate pure technical efficiency has contributed negatively. Thus, we find that the primary driver of aggregate productivity growth over the period 1990 to 2023 is the improvement in aggregate scale efficiency.

Table 3: Estimates of Unweighted Productivity Change and its Components Across 84 Countries/Regions

Period	$\exp(\widehat{\mathbb{M}}_n)$	$\exp(\widetilde{\mathbb{M}}_n)$	$\exp(\widehat{\mathbb{E}}_{1,n})$	$\exp(\widetilde{\mathbb{E}}_{1,n})$	$\exp(\widehat{\mathbb{E}}_{2,n})$	$\exp(\widetilde{\mathbb{E}}_{2,n})$	$\exp(\widehat{\mathbb{T}}_{1,n})$	$\exp(\widetilde{\mathbb{T}}_{1,n})$	$\exp(\widehat{\mathbb{S}}_{1,n})$	$\exp(\widetilde{\mathbb{S}}_{1,n})$
1990–1995	0.925	0.961	1.060	1.083**	1.031	1.097***	0.873	0.888***	1.028	0.987
1995–2000	1.039	1.050**	1.055	1.074***	1.040	1.039*	0.985	0.977	1.014	1.034***
2000–2005	1.072	1.086***	1.027	1.097***	1.031	1.074**	1.045	0.989	0.996	1.022
2005–2010	0.989	1.002	0.992	0.935**	1.009	0.950*	0.997	1.072***	0.984	0.984
2010–2015	0.989	0.994	1.096	1.255***	1.034	1.110***	0.903	0.792***	1.060	1.131***
2015–2020	0.920	0.920***	0.955	0.940**	0.935	0.913***	0.963	0.979**	1.020	1.030***
2020–2023	1.031	1.035***	0.924	0.852***	0.949	0.889***	1.117	1.215***	0.974	0.958***
1990–2023	1.157	1.461***	1.100	1.192***	1.023	1.064	1.052	1.225***	1.076	1.120***

NOTE: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (deviations from 1). For notational simplicity, we denote $\widetilde{\Gamma}_n = \widehat{\Gamma}_n - \widehat{B}_{\Gamma,n,\kappa,B}$, where $\Gamma \in \{\mathbb{M}, \mathbb{E}_1, \mathbb{E}_2, \mathbb{T}_1, \mathbb{S}_1\}$, i.e., $\widetilde{\Gamma}_n$ denotes the bias-corrected estimate of Γ.

Third, a comparison of the weighted results presented in Table 2 and the unweighted results in Table 3 reveals substantial differences between the simple mean and weighted aggregate approaches to the MPI. The notations $\widehat{\mathbb{M}}_n$, $\widehat{\mathbb{E}}_{1,n}$, $\widehat{\mathbb{E}}_{2,n}$, $\widehat{\mathbb{T}}_{1,n}$, and $\widehat{\mathbb{S}}_{1,n}$ reported in Table 3 denote the changes in productivity, conical technical efficiency, pure technical efficiency, conical technology, and scale efficiency, respectively, analogous to their aggregate counterparts presented in Table 2.

For example, Table 3 shows aggregate productivity rose significantly for the full sample in 1995–2000, 2000–2005, and 2020–2023. However, Table 2 indicates that there is no significant change in the aggregate productivity over these three periods. This comparison highlights the importance of considering weights when studying aggregate productivity change evolution for a group.

Moreover, while both tables show aggregate productivity for the full sample increased by approximately 50% from 1990 to 2023, Table 3 suggests that the growth during this period is attributable to improvements in both conical technical efficiency (1.192***) and technological progress (1.225***).

In contrast, Table 2 indicates that the aggregate productivity growth is primarily driven by gains in the aggregate conical technical efficiency (1.509**), while the aggregate conical technological change appears to be insignificant (0.987). These differences highlight the necessity of conducting statistical inference for both unweighted and weighted aggregate MPI measures to gain a more comprehensive understanding of productivity changes and their underlying sources within a group.

7 Conclusions

The Malmquist productivity index has become a standard instrument for quantifying productivity dynamics over time. A recent study by Pham et al. (2024) extends this strand of literature by formalizing statistical inference for the aggregate (weighted) MPI computed across a group of firms. Extending this line of research, the present paper develops a comprehensive inferential framework for the aggregated components of MPI-based productivity decomposition, which complements the unweighted methods developed by Simar and Wilson (2019).

Specifically, we derive central limit theorems for the weighted sources of the aggregate productivity change, which provide a rigorous basis for hypothesis testing and for constructing confidence intervals for the aggregate technological change, aggregate efficiency change, and other related components. This extension enables formal statistical assessment not only of overall productivity growth but also of its underlying sources.

To demonstrate the empirical relevance of our theoretical framework, we apply the proposed methods to data from the Penn World Table and investigate the sources of productivity change for 84 countries/regions. The empirical results reveal noticeable differences between the estimates obtained before and after bias correction for both the aggregate MPI and its components.

Furthermore, although both weighted and unweighted approaches indicate that the aggregate productivity rose by approximately 50% for the full sample over the period 1990–2023, substantial differences emerge between the unweighted and weighted sources of productivity change. In particular, the unweighted approach attributes the aggregate productivity growth to improvements in both conical technical efficiency and technological progress.

In contrast, the weighted approach suggests that the aggregate productivity growth is primarily driven by gains in conical technical efficiency, while technological progress exhibits no statistically significant change. These findings highlight the importance of considering both unweighted and weighted MPI measures to fully understand productivity dynamics and their underlying sources within a group.

Future research could extend the analytical framework developed in this paper to other productivity indices and their associated decompositions, such as the Hicks–Moorsteen productivity index proposed by Diewert (1992) and Bjurek (1996), and the Malmquist–Luenberger productivity index proposed by Chung et al. (1997). Establishing analogous inferential results for these alternative measures would further broaden the applicability of the proposed methodology.

Another promising direction is to incorporate undesirable outputs into the analysis and to develop corresponding decomposition frameworks for measuring and analyzing environmental productivity growth by utilizing the theoretical results for environmental efficiency proposed by Simar et al. (2025a). Deriving rigorous statistical inference procedures for environmental productivity change and its underlying sources would provide a formal basis for evaluating sustainability-oriented performance dynamics.

Declarations

Acknowledgement

Valentin Zelenyuk acknowledges the financial support from The University of Queensland. Shirong Zhao acknowledges the financial support from National Natural Science Foundation of China (grant number 72401056). The authors used AI tools for proofreading and language editing to improve the quality of the original (non-native) English writing, in a manner similar to professional proofreading and editing services. All results and interpretations remain entirely the authors' own.

Conflict of interest

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

References

Badunenko, O. and D. Romero-Ávila (2013), Financial development and the sources of growth and convergence, *International Economic Review* 54, 629–663.

Bjurek, H. (1996), The Malmquist total factor productivity index, *The Scandinavian Journal of Economics* 98, 303–313.

Caves, D. W., L. R. Christensen, and W. E. Diewert (1982), The economic theory of index numbers and the measurement of input, output and productivity, *Econometrica* 50, 1393–1414.

Chung, Y. H., R. Färe, and S. Grosskopf (1997), Productivity and undesirable outputs: A directional distance function approach, *Journal of Environmental Management* 51, 229–240.

Diewert, W. E. (1992), Fisher ideal output, input, and productivity indexes revisited, *Journal of Productivity Analysis* 3, 211–248.

Färe, R., S. Grosskopf, B. Lindgren, and P. Roos (1992), Productivity changes in Swedish pharmacies 1980–1989: A non-parametric Malmquist approach, *Journal of Productivity Analysis* 3, 85–101.

— (1994a), Productivity developments in Swedish hospitals: A Malmquist output approach, in A. Charnes, W. W. Cooper, A. Y. Lewin, and L. M. Seiford, eds., *Data Envelopment Analysis: Theory, Methodology and Applications*, Dordrecht: Kluwer Academic Publishers, pp. 253–272.

Färe, R., S. Grosskopf, and M. Norris (1997), Productivity growth, technical progress, and efficiency change in industrialized countries: reply, *The American Economic Review* 87, 1040–1044.

Färe, R., S. Grosskopf, M. Norris, and Z. Zhang (1994b), Productivity growth, technical progress, and efficiency change in industrialized countries, *American Economic Review* 84, 66–83.

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

Gilbert, A. and P. W. Wilson (1998), Effects of deregulation on the productivity of korean banks, *Journal of Economics and Business* 50, 133–155.

Henderson, D. J. and R. R. Russell (2005), Human capital and convergence: A production-frontier approach, *International Economic Review* 46, 1167–1205.

Kneip, A., B. Park, and L. Simar (1998), A note on the convergence of nonparametric DEA efficiency measures, *Econometric Theory* 14, 783–793.

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

— (2016), Testing hypotheses in nonparametric models of production, *Journal of Business and Economic Statistics* 34, 435–456.

— (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.

Kumar, S. and R. R. Russell (2002), Technological change, technological catch-up, and capital deepening: relative contributions to growth and convergence, *American Economic Review* 92, 527–548.

Ma, Z., J. Yin, and R. Su (2026), Efficiency bias correction and robustness analysis of the global Malmquist index approach, *European Journal of Operational Research* Forthcoming.

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.

Pastor, J. T., C. K. Lovell, and J. Aparicio (2020), Defining a new graph inefficiency measure for the proportional directional distance function and introducing a new Malmquist productivity index, *European Journal of Operational Research* 281, 222–230.

Pham, M. D., L. Simar, and V. Zelenyuk (2024), Statistical inference for aggregation of Malmquist productivity indices, *Operations Research* 72, 1615–1629.

Ramakrishna, G., M. Ramulu, and B. P. Kumar (2016), Financial crisis and the performance of commercial banks: Indian experience, *Journal of International Economics* 7, 35.

Ray, S. C. and E. Desli (1997), Productivity growth, technical progress, and efficiency change in industrialized countries: Comment, *American Economic Review* 87, 1033–1039.

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

Simar, L. and P. W. Wilson (1998a), Productivity growth in industrialized countries. Discussion Paper #9810, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

— (1998b), Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models, *Management Science* 44, 49–61.

— (2011), Inference by the m out of n bootstrap in nonparametric frontier models, *Journal of Productivity Analysis* 36, 33–53.

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

— (2023), Another look at productivity growth in industrialized countries, *Journal of Productivity Analysis* 60, 257–272.

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

Simar, L., V. Zelenyuk, and S. Zhao (2025a), Central limit theorems for directional distance functions with and without undesirable outputs. Unpublished working paper, Université catholique de Louvain, Institute of Statistics, Biostatistics and Actuarial Sciences (ISBA), Available at: `https://dial.uclouvain.be/pr/boreal/en/object/boreal%3A285841/datastream/PDF_01/view`.

— (2025b), Statistical inference for Hicks–Moorsteen productivity indices, *Annals of Operations Research* 351, 1675–1703.

Villa, G. and S. Lozano (2025), A global Malmquist productivity index of athletics performance in olympic games, *European Journal of Operational Research* Forthcoming.

Wheelock, D. C. and P. W. Wilson (1999), Technical progress, inefficiency, and productivity change in U. S. banking, 1984–1993, *Journal of Money, Credit, and Banking* 31, 212–234.

Wilson, P. W. and S. Zhao (2025), A non-parametric analysis of world productivity growth, 1990–2019, *Annals of Operations Research* 346, 2253–2285.

Zelenyuk, V. (2006), Aggregation of Malmquist productivity indexes, *European Journal of Operational Research* 174, 1076–1086.

— (2022), Aggregation of efficiency and productivity: from firm to sector and higher levels, in *Handbook of Production Economics*, Springer, pp. 1039–1079.

— (2023), Productivity analysis: roots, foundations, trends and perspectives, *Journal of Productivity Analysis* 60, 229–247.

— (2024), Aggregation in efficiency and productivity analysis: a brief review with new insights and justifications for constant returns to scale, *Journal of Productivity Analysis* 62, 321–334.

Zelenyuk, V. and S. Zhao (2025), Inference for DEA estimators of malmquist productivity indices: an overview, further improvements, and a guide for practitioners, *Macroeconomic Dynamics* 29, e81.