Centre for Efficiency and Productivity Analysis

Working Paper Series
No. WP04/2023

Further Improvements of Finite Sample Approximation of Central Limit Theorems for Weighted and Unweighted Malmquist Productivity Indices

Valentin Zelenyuk and Shirong Zhao

Date: May 2023

School of Economics
University of Queensland
St. Lucia, Qld. 4072
Australia

ISSN No. 1932 - 4398

Further Improvements of Finite Sample Approximation of Central Limit Theorems for Weighted and Unweighted Malmquist Productivity Indices

Valentin Zelenyuk*   Shirong Zhao†

May 2023

This is a draft not for circulation or citation without our permission

Abstract

Various methods recently have been proposed to further improve the finite sample performance of the developed central limit theorems (CLTs) for the simple mean and aggregate efficiency estimated via non-parametric frontier efficiency methods. We thoroughly investigate whether these methods are also effective to improve the finite sample performance for the recently developed CLTs for the simple mean and aggregate Malmquist Productivity Indices (MPIs). The extensive Monte-Carlo experiments confirmed that the method from Simar et al. (2023a) is useful for the simple mean and aggregate MPI in relatively small sample sizes (e.g., up to around 50, perhaps 100) and especially for large dimensions. Interestingly, we find that the better performance of the data sharpening method from Nguyen et al. (2022) observed in the context of efficiency is not obvious in the context of productivity. Finally, we use one well-known empirical data set to illustrate the differences across the existing methods to guide the practitioners.

Keywords: Malmquist Productivity Index, Non-parametric Efficiency Estimators, Data Envelopment Analysis, Inference

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

†Zhao: School of Finance, Dongbei University of Finance and Economics, Dalian, Liaoning 116025, China; email: shironz@dufe.edu.cn.

1 Introduction

The Malmquist productivity index (MPI) (Caves et al., 1982) is widely used to measure the productivity change of firms over time, and is typically estimated using Data Envelopment Analysis (DEA). For some examples of applications of MPI, see Färe et al. (1994), Ray and Desli (1997), Kumar and Russell (2002), Casu et al. (2013), Ramakrishna et al. (2016), Kevork et al. (2017), Pastor et al. (2020), Simar and Wilson (2022), to name a few.[1]

Recently, based on the seminal work of Kneip et al. (2015) for technical efficiency, Kneip et al. (2021) established statistical results for the individual MPI and geometric mean of MPIs, and also for these two measures in log terms. Based on Kneip et al. (2015), Simar and Zelenyuk (2018), and Kneip et al. (2021), Pham et al. (2023) further established the theoretical results for the weighted harmonic-type mean aggregation of MPIs, by taking the economic weight of each individual into account. These recent theoretical frameworks of MPI established by Kneip et al. (2021) and Pham et al. (2023), have enabled applied researchers for the first time to make theoretically well-grounded statistical inferences on the DEA-estimated productivity changes for a group of firms from various interesting economic questions.

However, it is observed from the Monte-Carlo (MC) simulations that for the simple mean of log MPIs (as evident in our MC results in Section 5) and the aggregate of log MPIs (Pham et al., 2023), the estimated confidence intervals based on the developed CLTs typically under-cover the true values in relatively small sample sizes, especially in large dimensions (measured by the total number of inputs and outputs). This under-covering phenomenon is also well observed in the other related non-parametric frontier efficiency estimators, such as the simple mean efficiency (Kneip et al., 2015) and the aggregate efficiency (Simar and Zelenyuk, 2018). This under-covering phenomenon mainly comes from the remaining bias in the estimation of the first and second moments. It is also amplified by the well known "curse of dimensionality" problem that the non-parametric methods typically suffer from, which states that the estimation errors are larger in finite sample sizes and in large dimensions, due to the slow convergence rates of the non-parametric methods.

To improve the finite sample approximation of these non-parametric frontier efficiency

[1] For a comprehensive review, see Ch4 and Ch7 of Sickles and Zelenyuk (2019).

estimators, Simar and Zelenyuk (2020), Nguyen et al. (2022), Simar et al. (2023a), and Simar et al. (2023b) have proposed various improving methods for the simple mean and aggregate (both input-oriented and output-oriented) efficiency. Their MC results suggest that the data sharpening method in Nguyen et al. (2022) and the method using bias-corrected efficiency estimates to obtain variance estimates in Simar et al. (2023a) are particularly useful in terms of larger coverages of the true values while also maintaining the developed central limit theorems for the simple mean and aggregate efficiency.

In this paper, our main objective is to examine whether the improving methods in Nguyen et al. (2022) and Simar et al. (2023a) developed specifically for the context of the simple mean and aggregate efficiency, are also useful to improve the finite sample approximation of CLTs for the simple mean and aggregate MPI, established by Kneip et al. (2021) and Pham et al. (2023), respectively. Through extensive Monte-Carlo experiments, we find that the method adapted from Simar et al. (2023a) could provide a better performance for the simple mean and aggregate MPI for relatively small samples (e.g., up to around 50, perhaps 100) and after that the original methods from Kneip et al. (2021) and Pham et al. (2023) are recommended. Moreover, we find that the better performance of the data sharpening method (Nguyen et al., 2022) observed in the simple mean and aggregate efficiency by Nguyen et al. (2022), Simar et al. (2023a), and Simar et al. (2023b), is not obvious in the simple mean and aggregate MPI, as the MPI estimates involve the ratios of technical efficiency estimates and so the magnitude of the bias seems to be partially cancelled out in practice.

Finally, using the well-known Penn World Table data set from 1990 to 2019 as an example, we illustrate the differences of all these methods in the estimated standard deviations and the significant levels for the simple mean and aggregate MPI, for the entire sample, developed and developing countries, and for pairs of years at 5–year intervals and the overall period 1990–2019.

The rest of the paper is organized as follows. Section 2 briefly introduces the theoretical background on technical efficiency, individual MPI, the simple mean and aggregate MPI, and the estimators of these measures. Section 3 briefly summarizes the main theoretical results from Kneip et al. (2021) and Pham et al. (2023) for the simple mean and aggregate MPI, respectively. Section 4 discusses how to adapt the improving methods from Nguyen et al. (2022) and Simar et al. (2023a) to the context of the simple mean and aggregate MPI.

Section 5 performs extensive MC experiments to evaluate the effectiveness of these improving methods. In Section 6, we use one real empirical data set to illustrate the differences of these methods in the estimated standard deviations and the significance levels. Section 7 concludes.

2 The Theoretical Background

2.1 The Production Economics Model

Denote $x \in \mathbb{R}_+^p$ and $y \in \mathbb{R}_+^q$ column vectors as inputs and outputs, respectively. The typical production set is

$$\Psi^t = \{(x, y) \mid x \text{ can produce } y \text{ at time } t\}, \tag{2.1}$$

which describes the set of physically attainable points (x, y) in the relevant input-output space at time t. We impose the common regularity assumptions on the production set Ψ^t, which are described in Appendix A in Pham et al. (2023). The upper boundary of Ψ^t, is called the *technology* or *frontier*, which is defined as

$$\Psi^{t\partial} := \left\{(x, y) \mid (x, y) \in \Psi^t,\ (x/\gamma, \gamma y) \notin \Psi^t \text{ for any } \gamma \in (1, \infty)\right\}. \tag{2.2}$$

The technical efficiency for a particular firm (x, y) is measured by the distance between the point of (x, y) in Ψ^t and the technology $\Psi^{t\partial}$. Efficiency can be measured in various orientations or directions. The Farrell output-oriented efficiency measure (Farrell, 1957) is

$$\lambda(x, y \mid \Psi^t) := \sup\{\lambda \mid (x, \lambda y) \in \Psi^t\}, \tag{2.3}$$

which gives the maximal proportion by which all outputs can be increased, while holding the inputs and technology fixed. For simplicity, in this paper, we focus on the output-oriented MPI. The results in this paper can be extended to the other directions.

The conical closure of the production set Ψ^t is defined as

$$\mathcal{C}(\Psi^t) := \left\{(\widetilde{x}, \widetilde{y}) \mid \widetilde{x} = ax,\ \widetilde{y} = ay,\ \forall\ a \in \mathbb{R}_+^1, \forall\ (x, y) \in \Psi^t\right\}. \tag{2.4}$$

If $\Psi^{t\partial}$ exhibits globally constant returns to scale (CRS), then $\mathcal{C}(\Psi^t) = \Psi^t$; otherwise, $\Psi^t \subset \mathcal{C}(\Psi^t)$. The corresponding conical Farrell output efficiency measure is defined as

$$\lambda_C(x, y \mid \Psi^t) := \lambda(x, y \mid \mathcal{C}(\Psi^t)) = \sup\{\lambda \mid (x, \lambda y) \in \mathcal{C}(\Psi^t)\}. \tag{2.5}$$

2.2 The Simple Mean and Aggregate MPI

Now consider a sample $\mathcal{S}_n = \{(X_i^1, Y_i^1), (X_i^2, Y_i^2)\}_{i=1}^n$ of input-output combinations for n firms observed in periods $t = 1$ and 2. To simplify the notation, let $Z_i^1 = (X_i^1, Y_i^1)$, $Z_i^2 = (X_i^2, Y_i^2)$, $\mathcal{S}_n^1 = \{Z_i^1\}_{i=1}^n$ and $\mathcal{S}_n^2 = \{Z_i^2\}_{i=1}^n$. We assume each firm i has access to the technology $\Psi^{t\partial}$ in period t, although potentially not efficient with respect to this technology. Following Caves et al. (1982), the productivity change for firm i from period 1 to period 2 can be defined as

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

Clearly, $\mathcal{M}_i > 1, = 1, \text{or } < 1$, indicates the productivity for firm i has increased, remained unchanged or decreased from period 1 to period 2.

In addition to estimating the productivity change for individual firms, applied researchers often are interested in whether the productivity change for a group, such as the geometric means of the individual productivity change,

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

is significantly greater or less than 1. Note that $\mathcal{M}$ uses the equally weighted geometric mean. Now consider the log MPI for individual firm as

$$\begin{aligned} \log \mathcal{M}_i = -\frac{1}{2} \Big[&\log \lambda_C(Z_i^2 \mid \Psi^1) + \log \lambda_C(Z_i^2 \mid \Psi^2) \\ &- \log \lambda_C(Z_i^1 \mid \Psi^1) - \log \lambda_C(Z_i^1 \mid \Psi^2) \Big], \end{aligned} \tag{2.8}$$

and denote the mean value as

$$\mu_{\mathcal{M}} := E(\log \mathcal{M}_i). \tag{2.9}$$

The log MPI for a group of firms is

$$\mu_{\mathcal{M},n} := \log \mathcal{M} = \frac{1}{n} \sum_{i=1}^n \log \mathcal{M}_i, \tag{2.10}$$

which is an estimate of $\mu_{\mathcal{M}}$ if the true value for individual MPI, $\mathcal{M}_i$, is known.

Another alternative is to take individual economic importance (such as the revenues) into account and consider the aggregate MPI (Zelenyuk, 2006), defined as

$$\overline{M} := \left(\frac{\sum_{i=1}^n \beta_i^2 \lambda_C(Z_i^2 \mid \Psi^1)}{\sum_{i=1}^n \beta_i^1 \lambda_C(Z_i^1 \mid \Psi^1)} \times \frac{\sum_{i=1}^n \beta_i^2 \lambda_C(Z_i^2 \mid \Psi^2)}{\sum_{i=1}^n \beta_i^1 \lambda_C(Z_i^1 \mid \Psi^2)} \right)^{-1/2}, \tag{2.11}$$

where

$$\beta_i^t = \frac{w^t Y_i^t}{\sum_{i=1}^n w^t Y_i^t}, \tag{2.12}$$

is the revenue weight for firm i at time t, and $w^t \in \mathbb{R}^q_{++}$ is the row vector of output prices, assumed to be the same for different firms in the same period t.

Now, consider the log version of $\overline{M}$ as

$$\begin{aligned}
\xi_n = \log \overline{M} = &- \frac{1}{2}\Bigg[\log\Big(\sum_{i=1}^n \beta_i^2 \lambda_C(Z_i^2 \mid \Psi^1)\Big) + \log\Big(\sum_{i=1}^n \beta_i^2 \lambda_C(Z_i^2 \mid \Psi^2)\Big) \\
&- \log\Big(\sum_{i=1}^n \beta_i^1 \lambda_C(Z_i^1 \mid \Psi^1)\Big) - \log\Big(\sum_{i=1}^n \beta_i^1 \lambda_C(Z_i^1 \mid \Psi^2)\Big)\Bigg] \\
= &- \frac{1}{2}\Bigg[\log\Big(\sum_{i=1}^n \lambda_C(Z_i^2 \mid \Psi^1) w^2 Y_i^2\Big) + \log\Big(\sum_{i=1}^n \lambda_C(Z_i^2 \mid \Psi^2) w^2 Y_i^2\Big) \\
&- \log\Big(\sum_{i=1}^n \lambda_C(Z_i^1 \mid \Psi^1) w^1 Y_i^1\Big) - \log\Big(\sum_{i=1}^n \lambda_C(Z_i^1 \mid \Psi^2) w^1 Y_i^1\Big)\Bigg] \\
&+ \log\Big(\sum_{i=1}^n w^2 Y_i^2\Big) - \log\Big(\sum_{i=1}^n w^1 Y_i^1\Big).
\end{aligned} \tag{2.13}$$

As shown by Pham et al. (2023), ξ_n is an consistent estimate of

$$\xi = -\frac{1}{2}(\log \mu_1 + \log \mu_2 - \log \mu_3 - \log \mu_4) + \log \mu_5 - \log \mu_6, \tag{2.14}$$

where $\mu_s = E(U_{s,i})$, $s = 1, 2, \ldots, 6$ and

$$\begin{aligned}
U_{1,i} &= \lambda_C(Z_i^2 \mid \Psi^1) w^2 Y_i^2, \\
U_{2,i} &= \lambda_C(Z_i^2 \mid \Psi^2) w^2 Y_i^2, \\
U_{3,i} &= \lambda_C(Z_i^1 \mid \Psi^1) w^1 Y_i^1, \\
U_{4,i} &= \lambda_C(Z_i^1 \mid \Psi^2) w^1 Y_i^1, \\
U_{5,i} &= w^2 Y_i^2, \\
U_{6,i} &= w^1 Y_i^1.
\end{aligned} \tag{2.15}$$

However, all the above quantities of $\lambda_C(Z_i^2 \mid \Psi^1)$, $\lambda_C(Z_i^2 \mid \Psi^2)$, $\lambda_C(Z_i^1 \mid \Psi^1)$, and $\lambda_C(Z_i^1 \mid \Psi^2)$ are the so-called true quantities of interest, derived and based on economic theory reasoning, which are usually unobserved in practice, and must be estimated from the sample data, as discussed in the next sections.

2.3 DEA Estimators

In the empirical analysis, we do not observe $\lambda_C(Z_i^2 \mid \Psi^1)$, $\lambda_C(Z_i^2 \mid \Psi^2)$, $\lambda_C(Z_i^1 \mid \Psi^1)$, and $\lambda_C(Z_i^1 \mid \Psi^2)$, and thus we do not observe $\mu_{\mathcal{M},n}$ and ξ_n and hence they must be estimated from the sample data.

Given a random sample $\mathcal{S}_n$, the conical Farrell output efficiency $\lambda_C(x, y \mid \Psi^t)$ can be estimated by the DEA estimator as,

$$\widehat{\lambda}_C(x, y \mid \mathcal{S}_n^t) = \max_{\lambda, s_1, \dots, s_n} \Big\{ \lambda \mid \lambda y \le \sum_{i=1}^n s_i Y_i^t,\ x \ge \sum_{i=1}^n s_i X_i^t,\ \forall\ s_i \ge 0 \Big\}. \tag{2.16}$$

The simple mean MPI, $\mu_{\mathcal{M}}$, can then be estimated by

$$\widehat{\mu}_{\mathcal{M},n} = \frac{1}{n} \sum_{i=1}^n \log \widehat{\mathcal{M}}_i, \tag{2.17}$$

where

$$\begin{aligned} \log \widehat{\mathcal{M}}_i = -\frac{1}{2} \Big[& \log \widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^1) + \log \widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^2) \\ & - \log \widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^1) - \log \widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^2) \Big]. \end{aligned} \tag{2.18}$$

Similarly, the aggregate MPI, ξ, can be estimated by

$$\widehat{\xi}_n = -\frac{1}{2}(\log \widehat{\mu}_1 + \log \widehat{\mu}_2 - \log \widehat{\mu}_3 - \log \widehat{\mu}_4) + \log \widehat{\mu}_5 - \log \widehat{\mu}_6, \tag{2.19}$$

where $\widehat{\mu}_s = \frac{1}{n} \sum_{i=1}^n \widehat{U}_{s,i}$, $s = 1, 2, \dots, 6$ and

$$\begin{aligned} \widehat{U}_{1,i} &= \widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^1) w^2 Y_i^2, \\ \widehat{U}_{2,i} &= \widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^2) w^2 Y_i^2, \\ \widehat{U}_{3,i} &= \widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^1) w^1 Y_i^1, \\ \widehat{U}_{4,i} &= \widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^2) w^1 Y_i^1, \\ \widehat{U}_{5,i} &= w^2 Y_i^2, \\ \widehat{U}_{6,i} &= w^1 Y_i^1. \end{aligned} \tag{2.20}$$

3 Central Limit Theorems and Inferences

The statistical properties for the estimators $\widehat{\mu}_{\mathcal{M},n}$ and $\widehat{\xi}_n$ have been well established by Kneip et al. (2021) and Pham et al. (2023), respectively. In this section, we briefly summarize their main results, in order to adapt the improving methods in Nguyen et al. (2022) and Simar et al. (2023a) to the context of the simple mean and aggregate MPI.

3.1 CLTs and Inferences for the Simple Mean MPI

Notice that $\widehat{\mu}_{\mathcal{M},n}$ is a biased estimator of $\mu_{\mathcal{M}}$. The bias of $\widehat{\mu}_{\mathcal{M},n}$ comes from the bias of various efficiency estimates $\widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^1)$, $\widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^2)$, $\widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^1)$, and $\widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^2)$. According to Kneip et al. (2021), the bias of $\widehat{\mu}_{\mathcal{M},n}$ can be consistently estimated using the generalized jackknife method. The procedures described by Kneip et al. (2021) are as follows.

For each $k = 1, 2, \ldots, K$, where $K \ll \binom{n}{\lfloor n/2 \rfloor}$,[2] split the sample into two evenly sized subsamples (for simplicity, assuming n is even) $\mathcal{S}_{1,n/2,k}$ and $\mathcal{S}_{2,n/2,k}$, so that $\mathcal{S}_{1,n/2,k} \cap \mathcal{S}_{2,n/2,k} = \emptyset$ and $\mathcal{S}_{1,n/2,k} \cup \mathcal{S}_{2,n/2,k} = \mathcal{S}_n$. Further, for each $l = 1, 2$, let $\mathcal{S}_{l,n/2,k}^1$ and $\mathcal{S}_{l,n/2,k}^2$ be the observations in $\mathcal{S}_{l,n/2,k}$, split by periods 1 and 2, respectively. For each $l = 1, 2$, compute

$$\begin{aligned}\widehat{\mu}_{\mathcal{M},l,k} = \frac{2}{n} \sum_{(Z_i^1, Z_i^2) \in \mathcal{S}_{l,n/2,k}} -\frac{1}{2}\Big[&\log \widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_{l,n/2,k}^1) + \log \widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_{l,n/2,k}^2) \\ &- \log \widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_{l,n/2,k}^1) - \log \widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_{l,n/2,k}^2)\Big].\end{aligned} \tag{3.1}$$

Then compute

$$\widehat{\mu}_{\mathcal{M},n,k}^* = \frac{1}{2}(\widehat{\mu}_{\mathcal{M},1,k} + \widehat{\mu}_{\mathcal{M},2,k}). \tag{3.2}$$

Repeating the above process for K times (large enough), we end up with the estimate of the bias term for $\widehat{\mu}_{\mathcal{M},n}$ given by

$$\widehat{B}_{\mathcal{M},n,\kappa,K} = \frac{1}{K} \sum_{k=1}^{K} (2^\kappa - 1)^{-1} (\widehat{\mu}_{\mathcal{M},n,k}^* - \widehat{\mu}_{\mathcal{M},n}), \tag{3.3}$$

where $\kappa = 2/(p + q + 1)$, if the true technology is VRS and $\kappa = 2/(p + q)$ if it is CRS. After obtaining the estimate of the bias term for $\widehat{\mu}_{\mathcal{M},n}$, the key results from Kneip et al. (2021) can be summarized in the following theorem.

Theorem 1. *Under the appropriate set of assumptions described in Theorem 3.6 of Kneip et al. (2021), for $\kappa \geq 2/5$, we have*

$$\sqrt{n}\left(\widehat{\mu}_{\mathcal{M},n} - \widehat{B}_{\mathcal{M},n,\kappa,K} - \mu_{\mathcal{M}} + R_{n,\mathcal{M},\kappa}\right) \xrightarrow{\mathcal{L}} N(0, \sigma_{\mathcal{M}}^2), \tag{3.4}$$

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

$$\sqrt{n_\kappa}\left(\widehat{\mu}_{\mathcal{M},n_\kappa} - \widehat{B}_{\mathcal{M},n,\kappa,K} - \mu_{\mathcal{M}} + R_{n,\mathcal{M},\kappa}\right) \xrightarrow{\mathcal{L}} N(0, \sigma_{\mathcal{M}}^2), \tag{3.5}$$

[2] $\lfloor n/2 \rfloor$ denotes the largest integer that is less than or equal to $n/2$.

where $R_{n,\mathcal{M},\kappa} = o(n^{-\kappa})$ *and* $\widehat{\mu}_{\mathcal{M},n_\kappa}$ *is a random subsample version, with size* $n_\kappa = \lfloor n^{2\kappa} \rfloor < n$, *of* $\widehat{\mu}_{\mathcal{M},n}$. *Formally,*

$$\widehat{\mu}_{\mathcal{M},n_\kappa} = \frac{1}{n_\kappa} \sum_{\{i|(Z_i^1,Z_i^2)\in\mathcal{S}_{n_\kappa}\}} \log \widehat{\mathcal{M}}_i, \tag{3.6}$$

where $\mathcal{S}_{n_\kappa}$ *is a random subsample, with the sample size* n_κ, *from* $\mathcal{S}_n$.[3]

Further, $\sigma^2_{\mathcal{M}}$ is not observed and must be estimated. To estimate $\sigma^2_{\mathcal{M}}$, Kneip et al. (2021) suggest using the empirical version of $\sigma^2_{\mathcal{M}}$, i.e.,

$$\widehat{\sigma}^2_{\mathcal{M}} = \frac{1}{n} \sum_{i=1}^{n} (\log \widehat{\mathcal{M}}_i - \widehat{\mu}_{\mathcal{M},n})^2. \tag{3.7}$$

The asymptotic $100(1-\alpha)\%$ confidence intervals for $\mu_{\mathcal{M}}$ are then given by

$$\left[\widehat{\mu}_{\mathcal{M},n} - \widehat{B}_{\mathcal{M},n,\kappa,K} \;\pm\; \Phi^{-1}_{1-\alpha/2}\, \widehat{\sigma}_{\mathcal{M}}/\sqrt{n}\right], \tag{3.8}$$

and

$$\left[\widehat{\mu}_{\mathcal{M},n_\kappa} - \widehat{B}_{\mathcal{M},n,\kappa,K} \;\pm\; \Phi^{-1}_{1-\alpha/2}\, \widehat{\sigma}_{\mathcal{M}}/\sqrt{n_\kappa}\right], \tag{3.9}$$

for $\kappa \geq 2/5$ and $\kappa < 1/2$, respectively. As noted by Kneip et al. (2021), if $\kappa = 2/5$, both (3.8) and (3.9) are applicable, but (3.9) is recommended due to a smaller remainder term.

3.2 CLTs and Inferences for the Aggregate MPI

Similarly, $\widehat{\xi}_n$ is a biased estimator of ξ. The bias of $\widehat{\xi}_n$ also comes from the bias of various efficiency estimates $\widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^1)$, $\widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^2)$, $\widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^1)$, and $\widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^2)$. According to Pham et al. (2023), the bias of $\widehat{\xi}_n$ can be consistently estimated using the generalized jackknife method. The procedures described by Pham et al. (2023) are as follows.

For each $k = 1, 2, \ldots, K$, where $K \ll \binom{n}{\lfloor n/2 \rfloor}$, split the sample into two evenly sized subsamples (for simplicity, assuming n is even) $\mathcal{S}_{1,n/2,k}$ and $\mathcal{S}_{2,n/2,k}$, so that $\mathcal{S}_{1,n/2,k} \cap \mathcal{S}_{2,n/2,k} = \emptyset$ and $\mathcal{S}_{1,n/2,k} \cup \mathcal{S}_{2,n/2,k} = \mathcal{S}_n$. Further, for each $l = 1, 2$, let $\mathcal{S}^1_{l,n/2,k}$ and $\mathcal{S}^2_{l,n/2,k}$ be the

[3] Note that here, $\widehat{\mathcal{M}}_i$ as defined in (2.18) is computed for each i in the sub-sample, but relative to all the data points in $\mathcal{S}_n$ rather than $\mathcal{S}_{n_\kappa}$, and then averaged over all i in that subsample (of size n_κ) to obtain $\widehat{\mu}_{\mathcal{M},n_\kappa}$.

observations in $\mathcal{S}_{l,n/2,k}$, split by periods 1 and 2, respectively. For each $l = 1, 2$, compute

$$\begin{aligned}
\widehat{\mu}_{1,l,k} &= \frac{2}{n} \sum_{Z_i^2 \in \mathcal{S}^1_{l,n/2,k}} \widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}^1_{l,n/2,k}) w^2 Y_i^2, \\
\widehat{\mu}_{2,l,k} &= \frac{2}{n} \sum_{Z_i^2 \in \mathcal{S}^2_{l,n/2,k}} \widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}^2_{l,n/2,k}) w^2 Y_i^2, \\
\widehat{\mu}_{3,l,k} &= \frac{2}{n} \sum_{Z_i^1 \in \mathcal{S}^1_{l,n/2,k}} \widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}^1_{l,n/2,k}) w^1 Y_i^1, \\
\widehat{\mu}_{4,l,k} &= \frac{2}{n} \sum_{Z_i^1 \in \mathcal{S}^2_{l,n/2,k}} \widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}^2_{l,n/2,k}) w^1 Y_i^1,
\end{aligned} \tag{3.10}$$

and define

$$\widehat{\xi}_{l,k} = -\frac{1}{2}(\log \widehat{\mu}_{1,l,k} + \log \widehat{\mu}_{2,l,k} - \log \widehat{\mu}_{3,l,k} - \log \widehat{\mu}_{4,l,k}) + \log \widehat{\mu}_5 - \log \widehat{\mu}_6. \tag{3.11}$$

Then compute

$$\widehat{\xi}^*_{n,k} = \frac{1}{2}(\widehat{\xi}_{1,k} + \widehat{\xi}_{2,k}). \tag{3.12}$$

Repeating the above process for K times, we end up with the estimate of the bias term for $\widehat{\xi}_n$ given by

$$\widehat{B}_{\xi,n,\kappa,K} = \frac{1}{K} \sum_{k=1}^{K} (2^\kappa - 1)^{-1} (\widehat{\xi}^*_{n,k} - \widehat{\xi}_n). \tag{3.13}$$

After obtaining the estimate of the bias term for $\widehat{\xi}_n$, the key results from Pham et al. (2023) can be summarized in the following theorem.

Theorem 2. *Under the appropriate set of assumptions described in Theorem 8 of Pham et al. (2023), for $\kappa \geq 2/5$, we have*

$$\sqrt{n}\left(\widehat{\xi}_n - \widehat{B}_{\xi,n,\kappa,K} - \xi + R_{n,\xi,\kappa}\right) \xrightarrow{\mathcal{L}} N(0, \sigma^2_\xi), \tag{3.14}$$

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

$$\sqrt{n_\kappa}\left(\widehat{\xi}_{n_\kappa} - \widehat{B}_{\xi,n,\kappa,K} - \xi + R_{n,\xi,\kappa}\right) \xrightarrow{\mathcal{L}} N(0, \sigma^2_\xi), \tag{3.15}$$

where $R_{n,\xi,\kappa} = o(n^{-\kappa})$ and $\widehat{\xi}_{n_\kappa}$ is a random subsample version, with size $n_\kappa = \lfloor n^{2\kappa} \rfloor < n$, of $\widehat{\xi}_n$. Formally,

$$\widehat{\xi}_{n_\kappa} = -\frac{1}{2}(\log \widehat{\mu}_{1,n_\kappa} + \log \widehat{\mu}_{2,n_\kappa} - \log \widehat{\mu}_{3,n_\kappa} - \log \widehat{\mu}_{4,n_\kappa}) + \log \widehat{\mu}_{5,n_\kappa} - \log \widehat{\mu}_{6,n_\kappa}, \tag{3.16}$$

where

$$\widehat{\mu}_{s,n_\kappa} = \frac{1}{n_\kappa} \sum_{\{i|(Z_i^1, Z_i^2) \in \mathcal{S}_{n_\kappa}\}} \widehat{U}_{s,i}, \;\; s = 1, 2, \ldots, 6, \tag{3.17}$$

and where $\mathcal{S}_{n_\kappa}$ is a random subsample, with the sample size n_κ, from $\mathcal{S}_n$.[4]

Further, σ_ξ^2 is not observed and must be estimated. To estimate σ_ξ^2, Pham et al. (2023) suggest plugging the corresponding empirical estimates of these components in σ_ξ^2, i.e.,

$$\widehat{\sigma}_\xi^2 = [\nabla \widehat{\xi}_n]' \widehat{\Sigma} [\nabla \widehat{\xi}_n], \tag{3.18}$$

where $\nabla \widehat{\xi}_n$ is the column vector of the gradient of $\widehat{\xi}_n$ with respect to $\widehat{\mu}_s$. Formally, $\nabla \widehat{\xi}_n = [\frac{\partial \widehat{\xi}_n}{\widehat{\mu}_1}, \frac{\partial \widehat{\xi}_n}{\widehat{\mu}_2}, \frac{\partial \widehat{\xi}_n}{\widehat{\mu}_3}, \frac{\partial \widehat{\xi}_n}{\widehat{\mu}_4}, \frac{\partial \widehat{\xi}_n}{\widehat{\mu}_5}, \frac{\partial \widehat{\xi}_n}{\widehat{\mu}_6}]'$, where

$$\begin{aligned} &\frac{\partial \widehat{\xi}_n}{\widehat{\mu}_1} = -\frac{1}{2\widehat{\mu}_1}, \;\; \frac{\partial \widehat{\xi}_n}{\widehat{\mu}_2} = -\frac{1}{2\widehat{\mu}_2}, \;\; \frac{\partial \widehat{\xi}_n}{\widehat{\mu}_3} = \frac{1}{2\widehat{\mu}_3}, \\ &\frac{\partial \widehat{\xi}_n}{\widehat{\mu}_4} = \frac{1}{2\widehat{\mu}_4}, \;\; \frac{\partial \widehat{\xi}_n}{\widehat{\mu}_5} = \frac{1}{\widehat{\mu}_5}, \;\; \frac{\partial \widehat{\xi}_n}{\widehat{\mu}_6} = \frac{1}{\widehat{\mu}_6}. \end{aligned} \tag{3.19}$$

Further, $\widehat{\Sigma}$ is the covariance matrix of a column vector $[\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\mu}_3, \widehat{\mu}_4, \widehat{\mu}_5, \widehat{\mu}_6]'$, Formally, the (j,k)th element of $\widehat{\Sigma}$ is

$$\widehat{\Sigma}_{j,k} = \frac{1}{n} \sum_{i=1}^{n} (\widehat{U}_{j,i} - \widehat{\mu}_j)(\widehat{U}_{k,i} - \widehat{\mu}_k). \tag{3.20}$$

The asymptotically $100(1-\alpha)\%$ confidence intervals for ξ are given by

$$\left[\widehat{\xi}_n - \widehat{B}_{\xi,n,\kappa,K} \;\pm\; \Phi^{-1}_{1-\alpha/2} \; \widehat{\sigma}_\xi / \sqrt{n} \right], \tag{3.21}$$

and

$$\left[\widehat{\xi}_{n_\kappa} - \widehat{B}_{\xi,n,\kappa,K} \;\pm\; \Phi^{-1}_{1-\alpha/2} \; \widehat{\sigma}_\xi / \sqrt{n_\kappa} \right], \tag{3.22}$$

for $\kappa \geq 2/5$ and $\kappa < 1/2$, respectively. In the remark 2 of Pham et al. (2023), if $\kappa = 2/5$, both (3.21) and (3.22) are applicable, but (3.22) is recommended due to a smaller remainder term.

[4] Note that here, $\widehat{U}_{s,i}$ as defined in (2.20) is computed for each i in the sub-sample, but relative to all the data points in $\mathcal{S}_n$ rather than $\mathcal{S}_{n_\kappa}$, and then averaged over all i in that subsample (of size n_κ) to obtain $\widehat{\mu}_{s,n_\kappa}$.

4 Further Improvements of Finite Sample Approximation of CLTs

It is observed from the simulations that for the simple mean (Kneip et al., 2021) and aggregate MPI (Pham et al., 2023), the estimated confidence intervals based on the developed CLTs typically under-cover the true values in finite sample sizes and large dimensions. For example, Table EC.9 in Pham et al. (2023) presents the coverages of confidence intervals when $p = 4$, $q = 1$, $\delta = 0.04$, which shows that when the nominal coverage is 95%, the estimated confidence intervals based on (3.22) for $n = 10, 20, 50, 100$ is only $0.684, 0.835, 0.902, 0.932$, respectively.

This under-covering phenomenon is also observed in the simple mean (Kneip et al., 2015) and aggregate efficiency (Simar and Zelenyuk, 2018). Recently, Simar and Zelenyuk (2020), Nguyen et al. (2022), Simar et al. (2023a), and Simar et al. (2023b) propose various methods to improve the finite sample approximation for the simple mean and aggregate (input-oriented and output-oriented) efficiency. In this section, to improve the finite sample approximation of CLTs for the simple mean of MPI and aggregate (weighted mean of) MPI, we adapt the methods from Nguyen et al. (2022) and Simar et al. (2023a).

4.1 Improvements via Data Sharpening

Specifically, we adapt the idea of the data sharpening method in Nguyen et al. (2022) to the output-oriented MPI. First, we sharpen $Z_i^1 = (X_i^1, Y_i^1)$ as follows,

$$\widetilde{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^1) = \begin{cases} \widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^1), & \text{if } 1/\widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^1) < 1 - \tau, \\ \\ \widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^1)/\varepsilon_i, & \text{otherwise }, \end{cases} \tag{4.1}$$

where the sharpening parameter $\tau = n^{-\gamma}$, while ε_i is a random independent number drawn from a uniform distribution on the interval $[1 - \tau, 1]$, where we set $\gamma = \kappa$.[5] It can be shown that $\widetilde{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^1) = \widehat{\lambda}_C(X_i^1, \widetilde{Y}_i^1 \mid \mathcal{S}_n^1)$, where

$$\widetilde{Y}_i^1 = \begin{cases} Y_i^1, & \text{if } 1/\widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^1) < 1 - \tau, \\ \\ Y_i^1 \times \varepsilon_i, & \text{otherwise }. \end{cases} \tag{4.2}$$

[5] Simar et al. (2023b) show that the main asymptotic results hold for any $\gamma > \min(\kappa/2, 1/4)$, and their extensive MC results suggest that choosing γ equal to or near κ, usually provides the best coverage.

After the data sharpening, (X_i^1, Y_i^1) becomes $(X_i^1, \widetilde{Y}_i^1)$. Moreover, we have

$$\widetilde{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^2) = \widehat{\lambda}_C(X_i^1, \widetilde{Y}_i^1 \mid \mathcal{S}_n^2) = \begin{cases} \widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^2), & \text{if } 1/\widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^1) < 1-\tau, \\ \\ \widehat{\lambda}_C(Z_i^1 \mid \mathcal{S}_n^2)/\varepsilon_i, & \text{otherwise .} \end{cases} \tag{4.3}$$

Similarly, we do the data sharpening for $Z_i^2 = (X_i^2, Y_i^2)$ as follows,

$$\widetilde{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^2) = \begin{cases} \widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^2), & \text{if } 1/\widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^2) < 1-\tau, \\ \\ \widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^2)/\epsilon_i, & \text{otherwise ,} \end{cases} \tag{4.4}$$

where the sharpening parameter $\tau = n^{-\kappa}$, while ϵ_i is a random independent number drawn from a uniform distribution on the interval $[1-\tau, 1]$. It can be shown that $\widetilde{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^2) = \widehat{\lambda}_C(X_i^2, \widetilde{Y}_i^2 \mid \mathcal{S}_n^2)$, where

$$\widetilde{Y}_i^2 = \begin{cases} Y_i^2, & \text{if } 1/\widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^2) < 1-\tau, \\ \\ Y_i^2 \times \epsilon_i, & \text{otherwise .} \end{cases} \tag{4.5}$$

After the data sharpening, (X_i^2, Y_i^2) becomes $(X_i^2, \widetilde{Y}_i^2)$. Moreover, we have

$$\widetilde{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^1) = \widehat{\lambda}_C(X_i^2, \widetilde{Y}_i^2 \mid \mathcal{S}_n^1) = \begin{cases} \widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^1), & \text{if } 1/\widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^2) < 1-\tau, \\ \\ \widehat{\lambda}_C(Z_i^2 \mid \mathcal{S}_n^1)/\epsilon_i, & \text{otherwise .} \end{cases} \tag{4.6}$$

Combining together, after the data sharpening, the input-output pair for observation i, $(X_i^1, Y_i^1, X_i^2, Y_i^2)$ becomes $(X_i^1, \widetilde{Y}_i^1, X_i^2, \widetilde{Y}_i^2)$. We then use the sharpened sample $\{(X_i^1, \widetilde{Y}_i^1, X_i^2, \widetilde{Y}_i^2)\}_{i=1}^n$ to obtain the estimates of the simple mean and aggregate MPI as well as their confidence intervals.

4.2 Improvements via Bias-corrected Individual Efficiency Estimates

Recently, Simar et al. (2023a) propose using the bias-corrected efficiency estimate to obtain the variance estimator for the simple mean efficiency, which is further extended by Simar et al. (2023b) for the aggregate efficiency. As the simple mean and aggregate MPI is constructed using various technical efficiency, the method in Simar et al. (2023a) can also be extended to here.

First, it is worth noting that when we estimate the bias for the simple mean MPI estimate $\widehat{\mu}_{\mathcal{M},n}$ or aggregate MPI estimate $\widehat{\xi}_n$, we can also obtain the bias for each individual technical efficiency estimate $\widehat{\lambda}_C(Z_i^s \mid \mathcal{S}_n^t)$, where $s,t \in \{1,2\}$. Specifically, when splitting the sample $\mathcal{S}_n$ into $\mathcal{S}_{1,n/2,k}$ and $\mathcal{S}_{2,n/2,k}$, we know that the observation i with the input-output pair (Z_i^1, Z_i^2) must lie in either $\mathcal{S}_{1,n/2,k}$ or $\mathcal{S}_{2,n/2,k}$ with equal probability. Without loss of generality, we assume $(Z_i^1, Z_i^2) \in \mathcal{S}_{1,n/2,k}$. Then we compute

$$\widehat{B}^*_{i,s,t,k} = \widehat{\lambda}_C(Z_i^s \mid \mathcal{S}^t_{1,n/2,k}) - \widehat{\lambda}_C(Z_i^s \mid \mathcal{S}_n^t). \tag{4.7}$$

Repeating the above process K times, we end up with the estimate of the bias term for $\widehat{\lambda}_C(Z_i^s \mid \mathcal{S}_n^t)$ given by

$$\widehat{B}_{i,s,t} = \frac{1}{K}\sum_{k=1}^{K}(2^\kappa - 1)^{-1}(\widehat{B}^*_{i,s,t,k}). \tag{4.8}$$

Extending the idea of Simar et al. (2023a) to the simple mean and aggregate MPI, for the original variance estimator of simple mean and aggregate MPI expressed in equations (3.7) and (3.18), respectively, we propose replacing $\widehat{\lambda}_C(Z_i^s \mid \mathcal{S}_n^t)$ by $\widehat{\lambda}_C(Z_i^s \mid \mathcal{S}_n^t) - \widehat{B}_{i,s,t}$ at every place, where $s,t \in \{1,2\}$, and $\widehat{B}_{i,s,t}$ is the estimated individual bias for $\widehat{\lambda}_C(Z_i^s \mid \mathcal{S}_n^t)$ using the generalized jackknife method of Kneip et al. (2015), as discussed above.

5 Monte-Carlo Evidence

5.1 Details on Monte-Carlo Simulations

Our MC experiments closely follow that in Pham et al. (2023), and we briefly restate it here for the sake of being self-contained. Interested readers can see Appendix EC.3 in Pham et al. (2023) for more details. The true technology in the first period is

$$y_i^{1\partial} = \prod_{j=1}^{p}(x_{ij}^1 - 1)^{\beta_j}, \tag{5.1}$$

while the true technology in the second period is

$$y_i^{2\partial} = (1+\delta)\prod_{j=1}^{p}(x_{ij}^2 - 1)^{\beta_j+\delta}, \tag{5.2}$$

where δ controls the changes of the technology from period 1 to period 2. Denote $x_i^t = (x_{i1}^t, x_{i2}^t, \ldots, x_{ip}^t)$, for $t = 1, 2$ and $i = 1, \ldots, n$, then $(x_i^1, y_i^{1\partial})$ and $(x_i^2, y_i^{2\partial})$ are the points on the technology in (5.1) and (5.2), respectively.

The inputs and technical efficiency between the two periods are typically correlated. To account for the correlations of inputs, following Pham et al. (2023), for each $j = 1, \ldots, p$, we generate the observed inputs as

$$x_{ij}^1, x_{ij}^2, \overset{\text{iid}}{\sim} \text{Unif}(1, 10), \forall\, i = 1, \ldots, n, \tag{5.3}$$

where $corr(x_{ij}^1, x_{ij}^2) = 0.5$, for each $j = 1, \ldots, p$.[6] Similarly, to account for the correlations of efficiency, we generate the true efficiency as

$$\lambda(X_i^1, Y_i^1 \mid \Psi^1), \lambda(X_i^2, Y_i^2 \mid \Psi^2) \overset{\text{iid}}{\sim} 1 + |N(0, 0.25^2)|, \forall\, i = 1, \ldots, n, \tag{5.4}$$

where $corr(\lambda(X_i^1, Y_i^1 \mid \Psi^1), \lambda(X_i^2, Y_i^2 \mid \Psi^2)) \neq 0$ and $N(0, 0.25^2)$ is the normal distribution with the variance 0.25^2 so that $|N(0, 0.25^2)|$ is the half normal distribution. The observed outputs then can be computed as

$$y_i^1 = \prod_{j=1}^{p} (x_{ij}^1 - 1)^{\beta_j} / \lambda(X_i^1, Y_i^1 \mid \Psi^1), \tag{5.5}$$

and

$$y_i^2 = (1 + \delta) \prod_{j=1}^{p} (x_{ij}^2 - 1)^{\beta_j + \delta} / \lambda(X_i^2, Y_i^2 \mid \Psi^2), \tag{5.6}$$

i.e., we project the optimal points from the corresponding technologies to the production set to obtain a simulated sample $\mathcal{S}_n = \{(x_i^1, y_i^1, x_i^2, y_i^2)\}_{i=1}^n$.

[6] See EC.3 in Pham et al. (2023) for more details on how to generate the correlated numbers.

Table 1: The Values of β_j and w_j

p	1	2	3	4	5	7
β_1	0.5	0.3	0.1	0.1	0.05	0.025
β_2		0.4	0.2	0.15	0.1	0.05
β_3			0.3	0.2	0.15	0.075
β_4				0.25	0.2	0.1
β_5					0.25	0.125
β_6						0.15
β_7						0.175

We set $\delta = 0.04$, the output prices $w_y^1 = w_y^2 = 1$, and the values of β_j are presented in Table 1. The number of MC trials for each experiment consisting of (n, p, q) is $1,000$. Moreover, we consider both the simple mean and aggregate Malmquist productivity indices. The true values of the simple mean and aggregate MPI are computed by following the steps in EC.3.8 of Pham et al. (2023). Before presenting our MC results, to simplify the notation, for simple mean MPI, we denote,

- (i): Using the standard central limit theorems, i.e., the unobserved elements being replaced by their respective DEA estimates.
- (ii): Kneip et al. (2021).[7]
- (iii): Kneip et al. (2021)+Simar et al. (2023a).
- (iv): Kneip et al. (2021)+Nguyen et al. (2022).
- (v): Kneip et al. (2021)+Nguyen et al. (2022) +Simar et al. (2023a).

For aggregate MPI, we denote

- (i): Using the standard central limit theorems, i.e., the unobserved elements being replaced by their respective DEA estimates.

[7] When $p + q < 4$, Theorem B.2 in Kneip et al. (2021) is used; otherwise, Theorem B.3 in Kneip et al. (2021) is used.

- (ii): Pham et al. (2023).[8]

- (iii): Pham et al. (2023)+Simar et al. (2023a).

- (iv): Pham et al. (2023)+Nguyen et al. (2022).

- (v): Pham et al. (2023)+Nguyen et al. (2022) +Simar et al. (2023a).

5.2 Monte-Carlo Results

5.2.1 General Remarks

We notice that when the nominal coverage is 99%, the coverage of the estimated confidence intervals for both the simple mean and aggregate MPI using either of (ii)–(v) is generally close to 99%, and thus we only focus on the cases where the nominal coverage is 90% and 95%. Further, the results in this section are robust to the values of δ. The empirical coverages from the simulation results for $\delta = 0.02$ and $\delta = 0.10$ are reported in the separate Appendix B and Appendix C, respectively, which have a similar pattern as those for $\delta = 0.04$.

5.2.2 Main Results

Figures 1 and 2 present the results for the coverage of the estimated confidence intervals for the simple mean MPI in the cases of the 90% and 95% nominal coverage, respectively, while Figures 3 and 4 present similar results for aggregate MPI.[9] We notice that when the sample size increases, the estimated coverage of (ii) in Figures 1–4 shows a gradual improvement in the approximation of the respective nominal coverage across different dimensions, which supports the developed theories for the simple mean MPI in Kneip et al. (2021) and the aggregate MPI in Pham et al. (2023).

Comparing (iii) with (ii), across different nominal coverage, we see that the coverage using (iii) is larger than or equal to that using (ii) for the simple mean MPI. For the aggregate MPI, this is also true in general, except for the three cases $p = q = 1$, $n = 20, 50, 200$ with 90% nominal coverage, the one case $p = q = 1$, $n = 100$ with 95% nominal coverage, and

[8] More specifically, when $p + q < 4$, equation (62) in Theorem 8 in Pham et al. (2023) is used; otherwise, equation (63) in Theorem 8 in Pham et al. (2023) is used.

[9] Tables A.1 and A.2 in the separate Appendix A present the values for the coverage of estimated confidence intervals for the simple mean and aggregate MPI, respectively.

the one case $p = q = 1$, $n = 20$ with 99% nominal coverage.[10] However, the goal should not be just larger per se, but closer to the nominal level. The improvements (measured by the closeness to the nominal level) of (iii) over (ii) are mainly observed in high dimensions ($p \geq 2$) and relatively small sample sizes ($n \leq 50$). For example, for the simple mean MPI, when $p = 3$, $n = 20$ and the nominal coverage is 95%, the coverage using (iii) is 0.949, which is much closer to the 95% nominal coverage than 0.881% obtained using (ii). Moreover this difference is especially substantial in high dimensions. For example, for the simple mean MPI, holding $n = 20$ and the nominal coverage 95%, the difference of the coverage between (iii) and (ii) increases from 0.068 (0.949 − 0.881) to 0.121 (0.959 − 0.838) when the number of input increases from $p = 3$ to $p = 7$. However, it is observed that (iii) often starts overshooting after $n = 100$ (and sometimes from around $n = 50$), while at around $n = 100$, (ii) starts approximating the nominal levels relatively well and the additional improvements seem to not be needed, especially if they can add noise and overshoot with the nominal coverage.

The comparison between (iii) and (ii) suggests that the method in Simar et al. (2023a) generally could improve the finite sample approximation of the developed CLTs for simple mean and aggregate MPI in high dimensions ($p \geq 2$) and relatively small sample sizes ($n \leq 50$). To some extent, our results are consistent with Simar et al. (2023a) and Simar et al. (2023b) who also find the better performance of the method in Simar et al. (2023a) for improving the finite sample approximation of CLTs for simple mean efficiency and aggregate efficiency, respectively. However, our result is also different from the context of CLTs for efficiency aggregates, where the improvement was really needed, even for $n = 300$ and especially below that. This is likely due to the ratio nature of the measurement where the magnitude of the bias shrinks substantially.

Comparing (iv) with (ii), for both simple mean and aggregate MPI, the coverage using (iv) is generally smaller than that using (ii), except in high dimensions ($p \geq 4$) and small sample sizes ($n \leq 100$); in other words, the improvements of (iv) over (ii) are only observed in high dimensions and relatively small sample sizes. For example, for the simple mean MPI, when $p = 5$ and the nominal coverage is 95%, the difference of the coverage between (iv) and (ii) is 0.049 (0.912 − 0.863) when $n = 20$ and it decreases to −0.004 (0.942 − 0.946) when

[10] Note that the standard CLT (i) is also correct for the case $p = q = 1$.

Figure 1: Coverages of Estimated Confidence Intervals for the Simple Mean Malmquist Productivity Indices for the 90% Nominal Coverage

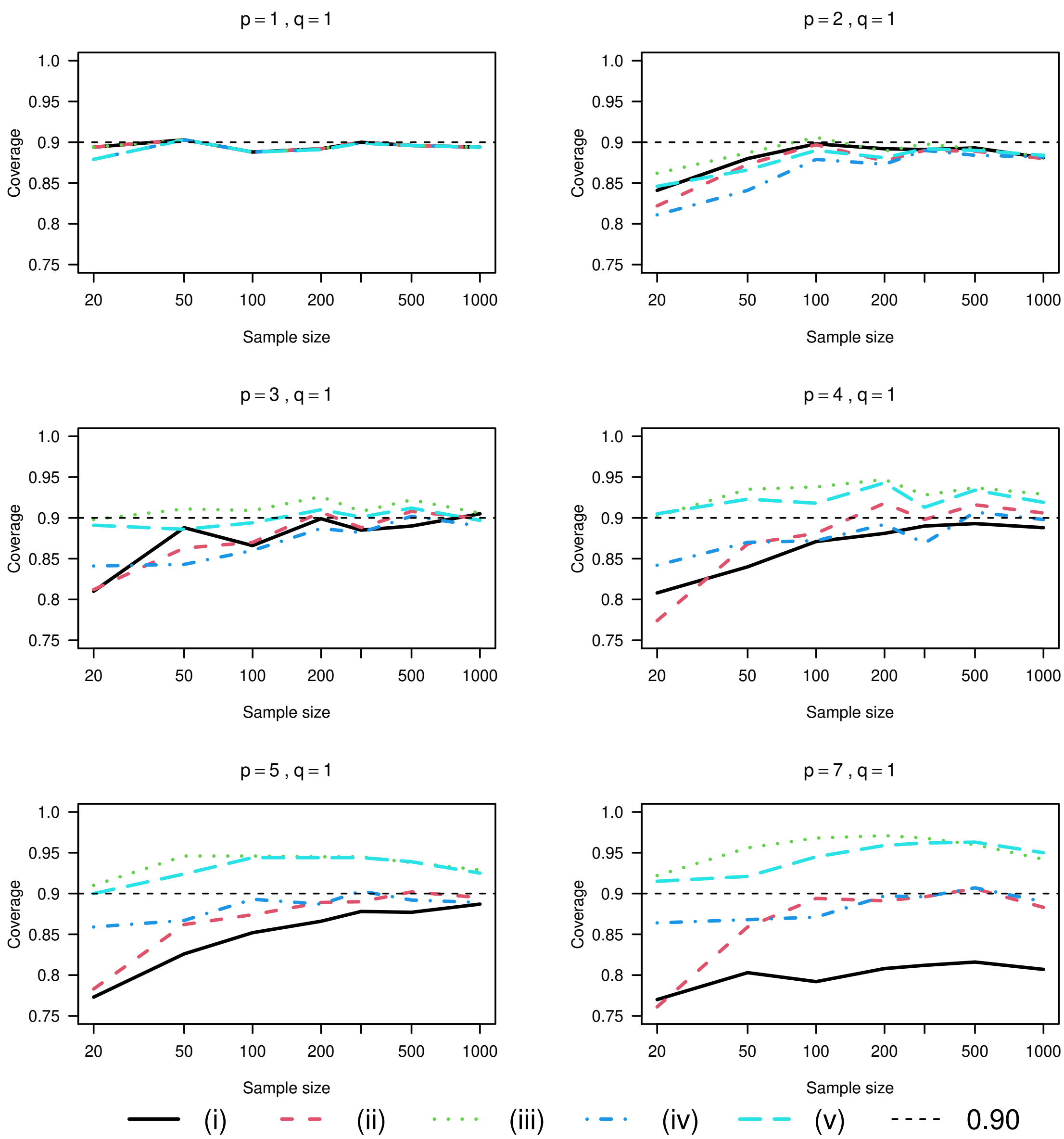


Figure 2: Coverages of Estimated Confidence Intervals for the Simple Mean Malmquist Productivity Indices for the 95% Nominal Coverage

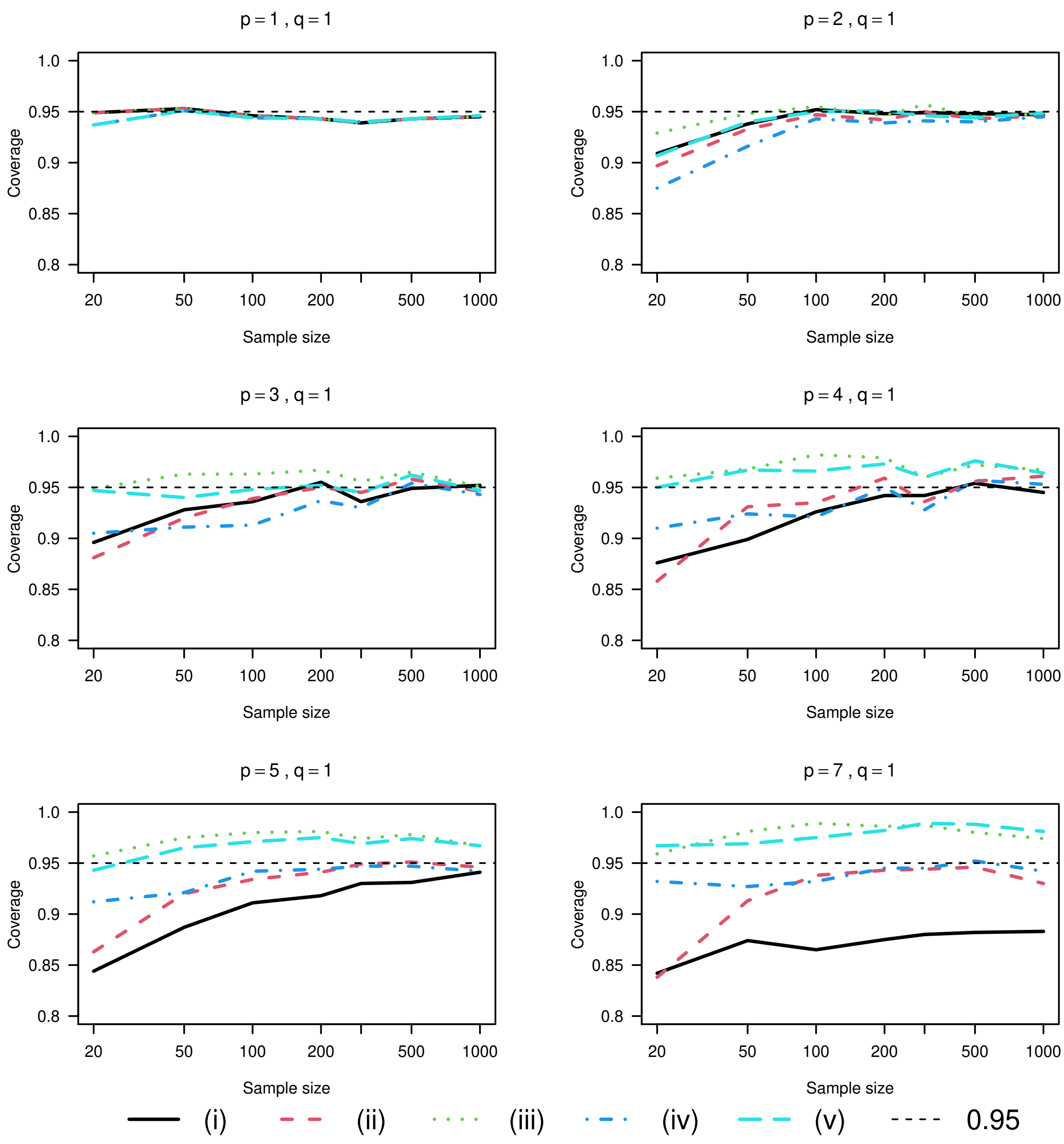


Figure 3: Coverages of Estimated Confidence Intervals for the Aggregate Malmquist Productivity Indices for the 90% Nominal Coverage

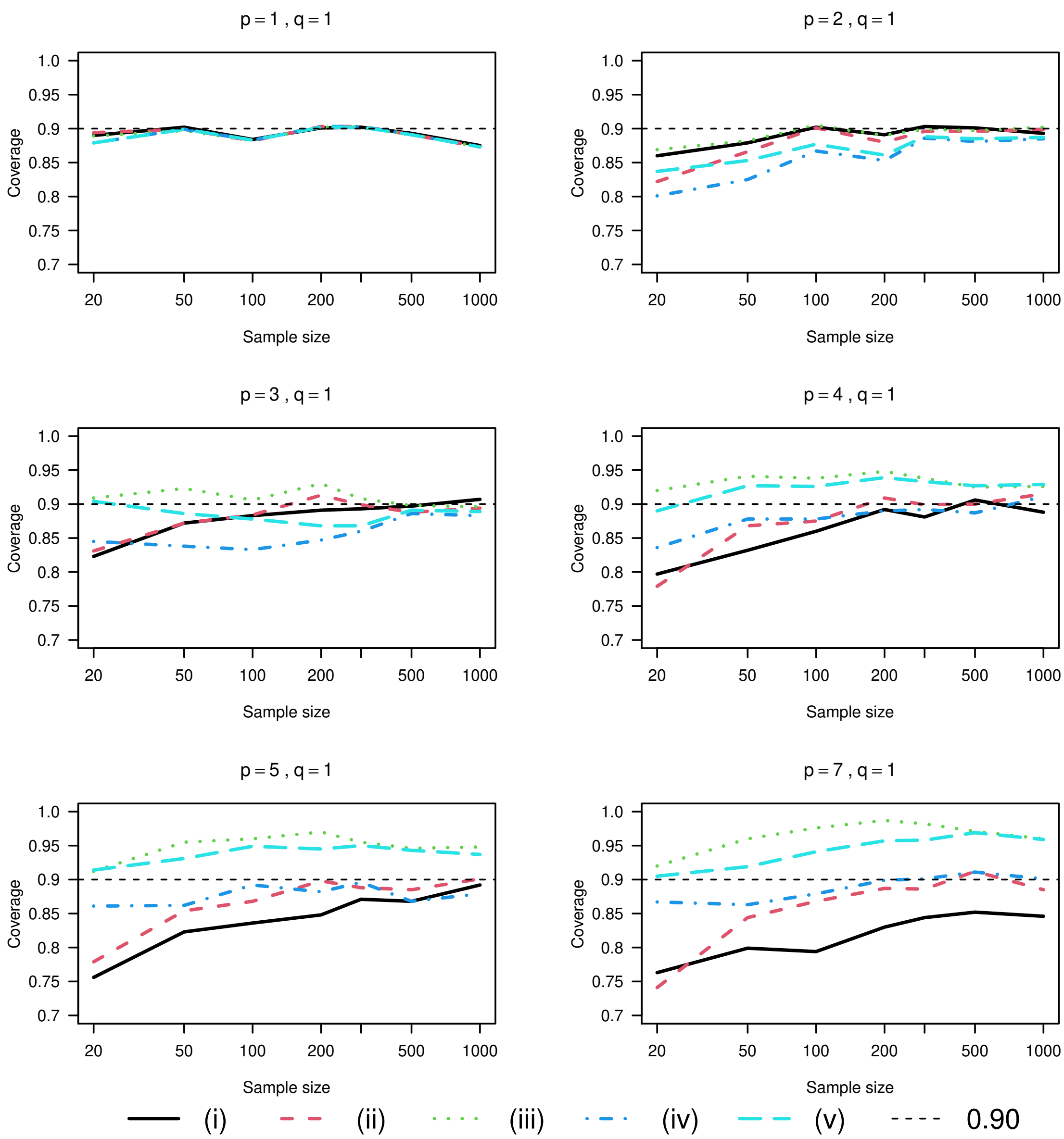
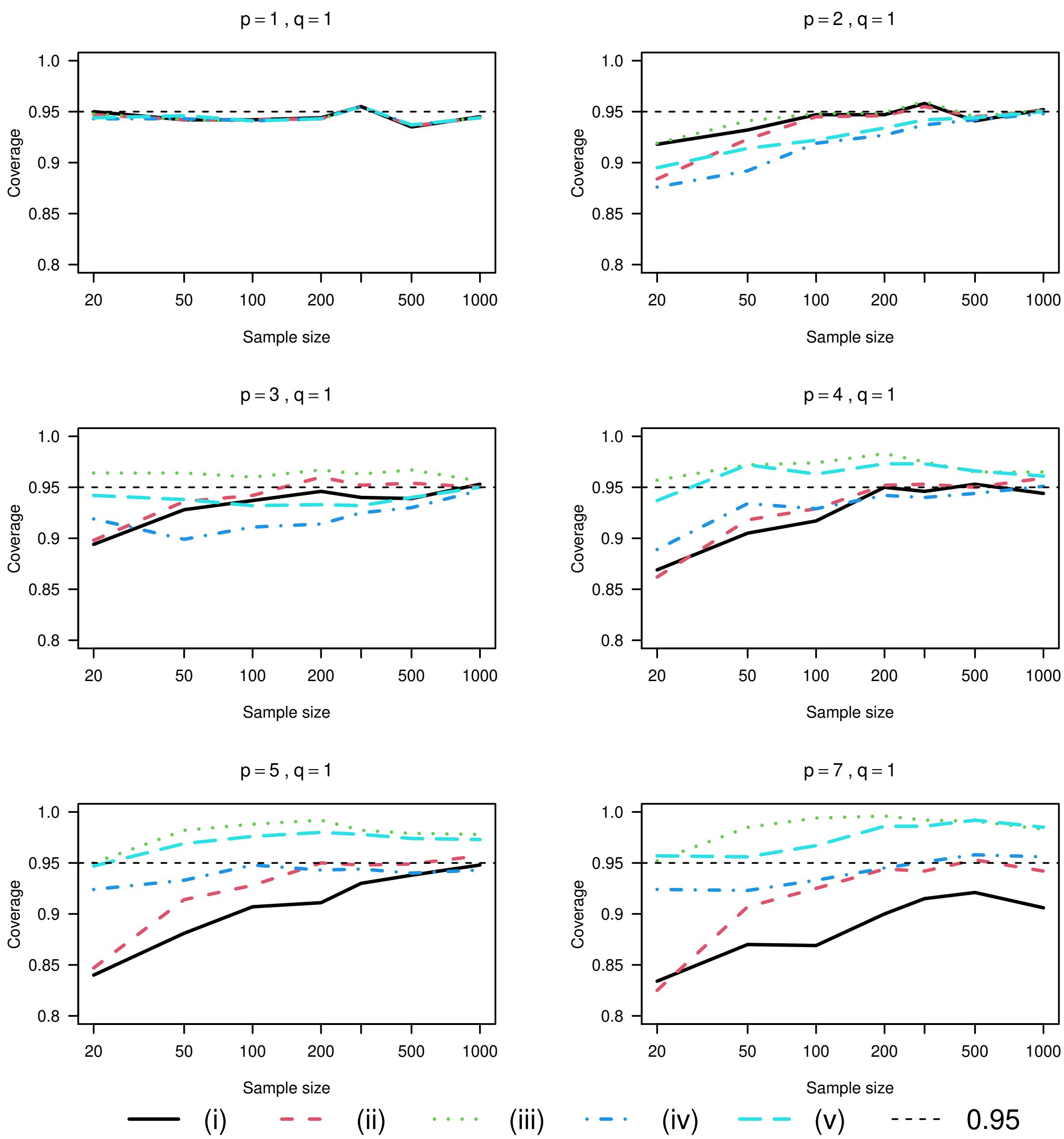


Figure 4: Coverages of Estimated Confidence Intervals for the Aggregate Malmquist Productivity Indices for the 95% Nominal Coverage



$n = 1000$. Our results here are different from Nguyen et al. (2022), Simar et al. (2023a), and Simar et al. (2023b) where they all find that the data sharpening method could improve the finite sample approximation of the central limit theorems for simple mean and aggregate efficiency; i.e., they all find the better performance of the data sharpening method over the original methods in Kneip et al. (2015) and Simar and Zelenyuk (2018). The reason might come from the fact that the MPI estimates involve ratios of technical efficiency estimates and so the magnitude (i.e., the constant part) of the bias seems to be partially cancelled out in practice.[11]

As both (iii) and (iv) have better performance over (ii) in high dimensions and relatively small sample sizes, next we compare these two methods. From Figures 1–4, we see that the improvements of (iii) are always more substantial than (iv) in all the high dimensions and relatively small sample sizes. Combining the results until now, the estimated coverage using (iii) seems to be much closer to the nominal coverage than those using (ii) and (iv) in high dimensions ($p \geq 2$) and relatively small sample sizes ($n \leq 50$), while at around $n = 100$, the method (ii) starts approximating the nominal levels relatively well.

Comparing (v) with (iv) again indicates the better performance of Simar et al. (2023a), as for both the simple mean and aggregate MPI, the coverage using (v) is larger than that using (iv), especially in high dimensions ($p \geq 2$) and small sample sizes ($n \leq 50$), (v) seems to often give slightly (and sometimes significantly) closer coverage than (iv). However, when comparing (v) and (iii), the former seems to often give slightly (and sometimes significantly) closer coverage than the latter, yet not always and the difference is often too small to justify the additional complexity. So, for simplicity reasons, (iii) might be preferred over (v).

To conclude for this section, both (iii) and (v) are very similar and help to improve the coverage for relatively small samples, such as around $n = 50$ and less, and sometimes up to around 100, but they also often start to overshoot after that (and sometimes from around 50). Fortunately, and unlike the context of efficiency measurement, already at around $n = 100$, the original methods from Kneip et al. (2021) and Pham et al. (2023) start approximating the nominal levels relatively well and the additional improvements seem to be not needed,

[11] For example, suppose the true efficiency scores are 1.4 and 1.5, but their DEA estimates are 1.3 and 1.4, respectively: the absolute value of the bias for both cases is 0.1, however if the interest is about the ratio of the true efficiency scores, i.e., 1.5/1.4=1.0714, it is still fairly well approximated by the ratio of their estimates, even though those estimates are biased (according to the same estimator). Indeed, 1.4/1.3=1.0769, implying the absolute value of the bias is only 0.0055.

especially if they can add noise and overshoot with the coverage (i.e., rejecting a hypothesis less than they should). This is indeed very different from the context of CLTs for efficiency aggregates, where the improvement was often needed even for $n = 300$ and especially below that. This is likely due to the ratio nature of the measurement in the MPI framework where the magnitude of the bias shrinks substantially. Hence, the bottom line conclusion is that the use of (iii) or (v) is advisable for relatively small samples (e.g., up to around 50, perhaps 100) and after that just use (ii), i.e., the original methods from Kneip et al. (2021) and Pham et al. (2023).

6 Empirical Illustrations

Our illustration closely follows that in Pham et al. (2023), which employs the widely used Penn World Table data (PWT 10.0, Feenstra et al., 2015) to study the MPI of countries/regions in the world from 1990 to 2019. For the related literature using PWT, see also Färe et al. (1994), Kumar and Russell (2002), and Badunenko et al. (2008, 2018).

Same as Pham et al. (2023), the production of countries is modelled using labor (*emp*) and capital stock (*cn*) to produce GDP (*rgdpo*). The output price is the same for all countries/regions. To illustrate the evolution of the simple mean and aggregate MPI from 1990 to 2019, we will use the same sub-set of 84 countries/regions as Badunenko et al. (2008) and Pham et al. (2023). Different from Pham et al. (2023), we present results for pairs of years at 5–year intervals and the overall period 1990–2019 in Table 2 for the simple mean MPI and in Table 3 for aggregate MPI. Moreover, same as Pham et al. (2023), we conduct the analysis for the entire sample, 27 developed countries, and 57 developing countries, separately.

First, from Tables 2 and 3, we can see that our results for the simple mean and aggregate MPI estimates from 1990 to 2019 are similar to those in Table 1 of Pham et al. (2023).[12] In terms of the estimates of the variance for the simple mean and aggregate MPI, we find that (iii) generally yields slightly larger estimates than (ii). For example, among the 21 cases in Table 2 (7 cases for each of the entire sample, developed and developing countries), the estimates of $\sigma_{\mathcal{M}}$ using (iii) are larger than that using (ii) for 19 cases; similarly, Table 3

[12] For example, we find that the bias-corrected simple mean MPI estimate ($\exp(\widehat{\mathcal{M}}_n - \widehat{B}_{\mathcal{M},n})$) is 1.0579, while it is 1.0640 in Pham et al. (2023). The small difference comes from the estimated bias for the simple mean and aggregate MPI, which uses the generalized jackknife method by randomly splitting the sample into two sub-samples many times.

Table 2: Estimation Results for the Simple Mean MPI of Countries/Regions

						— Estimate of $\sigma_{\mathcal{M}}$ —				— Significant Level —			
Year 1	Year 2	$\exp(\widehat{\mu}_{\mathcal{M},n})$	$\exp(\widehat{\mu}_{\mathcal{M},n} - \widehat{B}_{\mathcal{M},n,\kappa,K})$	$\widehat{\mu}_{\mathcal{M},n}$	$\widehat{\mu}_{\mathcal{M},n} - \widehat{B}_{\mathcal{M},n,\kappa,K}$	(ii)	(iii)	(iv)	(v)	(ii)	(iii)	(iv)	(v)
				— Entire Sample —									
1990	1995	0.8970	0.9121	-0.1087	-0.0921	0.3441	0.3540	0.3402	0.3497	**	**	**	**
1995	2000	1.0342	1.0371	0.0336	0.0364	0.1911	0.1910	0.1881	0.1877	*	*	*	*
2000	2005	1.0934	1.0912	0.0893	0.0873	0.2088	0.2141	0.2055	0.2104	***	***	***	***
2005	2010	0.9628	0.9786	-0.0380	-0.0216	0.2299	0.2444	0.2222	0.2373	–	–	–	–
2010	2015	0.9855	0.9848	-0.0146	-0.0153	0.1897	0.2041	0.1855	0.1999	–	–	–	–
2015	2019	1.0156	1.0157	0.0155	0.0156	0.1779	0.1953	0.1773	0.1944	–	–	–	–
1990	2019	1.0388	1.0579	0.0381	0.0563	0.4555	0.4678	0.4503	0.4617	–	–	–	–
				— Developed Countries —									
1990	1995	1.0532	1.0918	0.0519	0.0878	0.1067	0.1239	0.1337	0.1523	***	***	***	***
1995	2000	1.1706	1.1682	0.1576	0.1555	0.0785	0.0901	0.1086	0.1174	***	***	***	***
2000	2005	1.0720	1.0597	0.0695	0.0580	0.0881	0.0946	0.0939	0.0999	***	***	***	***
2005	2010	0.8654	0.8005	-0.1446	-0.2225	0.1284	0.1671	0.1148	0.1518	***	***	***	***
2010	2015	0.9295	0.8968	-0.0731	-0.1090	0.1122	0.1321	0.1019	0.1196	***	***	***	***
2015	2019	1.0403	1.0388	0.0395	0.0381	0.0393	0.0487	0.0583	0.0640	***	***	***	***
1990	2019	1.1284	1.0587	0.1208	0.0570	0.2046	0.2365	0.2085	0.2311	–	–	**	**
				— Developing Countries —									
1990	1995	0.8665	0.8908	-0.1433	-0.1157	0.4045	0.4249	0.3962	0.4167	**	**	**	**
1995	2000	0.9823	0.9922	-0.0179	-0.0078	0.1961	0.2002	0.1995	0.2036	–	–	–	–
2000	2005	1.1254	1.1309	0.1181	0.1230	0.2393	0.2482	0.2406	0.2498	***	***	***	***
2005	2010	1.0953	1.1647	0.0910	0.1525	0.2271	0.2248	0.2124	0.2090	***	***	***	***
2010	2015	1.0455	1.0691	0.0445	0.0668	0.2115	0.2198	0.2070	0.2152	**	**	**	**
2015	2019	1.0180	1.0211	0.0179	0.0208	0.2164	0.2324	0.2218	0.2377	–	–	–	–
1990	2019	1.1114	1.1772	0.1056	0.1632	0.5597	0.5861	0.5515	0.5767	**	**	**	**

NOTE: Statistical significance (difference from 1) for the bias-corrected estimate (i.e., $\exp(\widehat{\mu}_{\mathcal{M},n} - \widehat{B}_{\mathcal{M},n,\kappa,K})$) of the true mean of MPI at the ten, five, or one percent levels is denoted by one, two, or three asterisks, respectively, while "–" indicates insignificance at the ten percent level.

Table 3: Estimation Results for the Aggregate MPI of Countries/Regions

						— Estimate of σ_ξ —				— Significant Level —			
Year 1	Year 2	$\exp(\widehat{\xi}_n)$	$\exp(\widehat{\xi}_n - \widehat{B}_{\xi,n,\kappa,K})$	$\widehat{\xi}_n$	$\widehat{\xi}_n - \widehat{B}_{\xi,n,\kappa,K}$	(ii)	(iii)	(iv)	(v)	(ii)	(iii)	(iv)	(v)
				— Entire Sample —									
1990	1995	1.0057	1.0106	0.0057	0.0106	0.2361	0.2732	0.2275	0.2660	–	–	–	–
1995	2000	1.0453	1.0307	0.0443	0.0302	0.4584	0.5281	0.4003	0.4629	–	–	–	–
2000	2005	1.0202	0.9983	0.0200	-0.0017	0.2586	0.2963	0.2801	0.3146	–	–	–	–
2005	2010	0.9352	0.9651	-0.0670	-0.0355	0.1773	0.2714	0.1769	0.2712	*	–	*	–
2010	2015	0.9998	1.0199	-0.0002	0.0197	0.1526	0.2636	0.1726	0.2958	–	–	–	–
2015	2019	1.0034	1.0044	0.0034	0.0044	0.2098	0.1588	0.2314	0.1722	–	–	–	–
1990	2019	0.9850	0.9232	-0.0151	-0.0799	0.7554	0.6721	0.7907	0.7001	–	–	–	–
				— Developed Countries —									
1990	1995	1.0724	1.1113	0.0699	0.1056	0.0705	0.0855	0.1170	0.1106	***	***	***	***
1995	2000	1.1513	1.1656	0.1409	0.1532	0.0582	0.1146	0.0617	0.1099	***	***	***	***
2000	2005	1.0509	1.0563	0.0496	0.0548	0.0583	0.0546	0.0872	0.0717	***	***	***	***
2005	2010	0.8986	0.8751	-0.1069	-0.1334	0.1789	0.2139	0.1557	0.1767	***	***	***	***
2010	2015	0.9854	0.9744	-0.0147	-0.0259	0.1712	0.2063	0.2124	0.2426	–	–	–	–
2015	2019	1.0340	1.0312	0.0334	0.0307	0.0311	0.0410	0.1015	0.0961	***	***	***	***
1990	2019	1.1570	1.1583	0.1458	0.1469	0.3214	0.3618	0.3912	0.4262	**	**	***	***
				— Developing Countries —									
1990	1995	0.9389	0.9170	-0.0631	-0.0866	0.4814	0.5272	0.4787	0.5229	–	–	–	–
1995	2000	0.8930	0.8152	-0.1132	-0.2043	0.4151	0.4694	0.3994	0.4552	***	***	***	***
2000	2005	0.9963	0.9083	-0.0037	-0.0961	0.3676	0.4079	0.3672	0.4083	**	*	*	*
2005	2010	1.0530	1.1372	0.0516	0.1285	0.2378	0.2164	0.2362	0.2154	***	***	***	***
2010	2015	1.0510	1.1000	0.0498	0.0953	0.1428	0.2189	0.1436	0.2184	***	***	***	***
2015	2019	0.9968	0.9708	-0.0032	-0.0297	0.2733	0.2615	0.2687	0.2581	–	–	–	–
1990	2019	1.1297	1.0447	0.1219	0.0437	0.3803	0.3290	0.3663	0.3106	–	–	–	–

NOTE: Statistical significance (difference from 1) for the bias-corrected estimate (i.e., $\exp(\widehat{\xi}_n - \widehat{B}_{\xi,n,\kappa,K})$) of the true aggregate of MPI at the ten, five, or one percent levels is denoted by one, two, or three asterisks, respectively, while "–" indicates insignificance at the ten percent level.

shows that the estimates of σ_ξ using (iii) are larger than that using (ii) for 16 cases out of 21 cases. This result suggests that the estimates of confidence intervals based on (iii) generally will be slightly larger than that using (ii). This is consistent with our MC results which also suggested that (iii) has a better performance (in terms of covering the true values) than (ii) in relatively small sample sizes (e.g., up to around 50) and large dimensions. Similarly, comparing (v) and (iv), we find that (v) generally yields slightly larger estimates of the variance than (iv). Table 2 shows that the estimates of $\sigma_{\mathcal{M}}$ using (v) are larger than that using (iv) for 19 cases out of 21 cases and Table 3 shows that the estimates of σ_ξ using (v) are larger than that using (iv) for 13 cases out of 21 cases. This result suggests that the estimates of confidence intervals based on (v) generally will be slightly larger than that using (iv). Thus, our illustration confirms again our main takeaways from the MC results that the use of (iii) or (v) is advisable for relatively small samples (e.g., up to around 50, perhaps 100).

From Table 2, we see that the simple mean MPI for the entire sample is significantly different from 1 for 1990–2005; Further, this result is robust as the 95% CI constructed using either of (ii), (iii), (iv), and (v) does not contain 1, suggesting that the productivity growth of these 84 countries significantly decreased from 1990 to 1995, significantly increased from 1995 to 2005 and remained unchanged for the other remaining periods. For the developed countries, the simple mean MPI is significantly larger than 1 in the periods 1990–1995, 1995–2000, 2000–2005, and 2015–2019, while it is significantly smaller than 1 in the periods 2005–2010 and 2010–2015. This result is robust across different methods, suggesting that the productivity growth for the developed countries continued increasing from 1990 to 2005, then continued decreasing from 2005 to 2015 (possibly due to the global financial crisis) and increased again from 2015 to 2019. However, for the whole period 1990–2019, (ii) and (iii) suggest that productivity remained unchanged while (iv) and (v) suggest that productivity increased significantly. The productivity growth for the developing countries significantly decreased from 1990 to 1995, continued increasing from 2000 to 2015, and increased over the whole period 1990–2019. Thus, the results indicate that from 1990 to 2019 only developing countries achieved about 17.72% percent increase in productivity, while there is no evidence supporting the changes in productivity for the full sample of 84 countries and mixed evidence for the developed countries. Recall that these estimates ignore the economic weights of each country in the averaging, e.g., weighting the productivity estimate for the USA with the same weight as for any other country. To address this we estimate the aggregate MPI as per our discussions above and report the results in Table 3.

From Table 3, we see that the aggregate MPI for the entire sample is not significantly different from 1 in most of the considered periods, except for 2005–2010, where only (ii) and (iv) find that the productivity decreased at the 10% significance level. For the developed countries, the aggregate MPI is significantly larger than 1 in the periods 1990–1995, 1995–2000, 2000–2005, 2015–2019 and over the whole period 1990–2019, while it is significantly smaller than 1 from 2005 to 2010. This result is robust across different methods, suggesting that the productivity growth for the developed countries continued increasing from 1990 to 2005, decreased from 2005 to 2010 (when recall the global financial crisis occurred), increased again from 2015 to 2019, and also increased over the whole period 1990–2019. The productivity growth for the developing countries continued decreasing from 1995 to 2005, continued increasing from 2005 to 2015. Thus, the results in Table 3 indicate that only developed countries achieved about 15.83% percent increase in productivity, while productivity for the full sample of all countries and for the sample of developing countries remained unchanged. This result is different from the simple mean MPI presented in Table 2, where we find that only developing countries achieved about 17.72% percent increase in productivity, while productivity for the full sample of all countries remained unchanged and the productivity change for the developed countries is mixed.

To conclude this section, we see fairly different results between the weighted and non-weighted approaches of MPIs, illustrating the importance of deploying both approaches to check the changes of productivity growth. Moreover, our illustration confirms again our main takeaways from the MC results that the use of (iii) or (v) is advisable for relatively small samples (e.g., up to around 50).

7 Conclusions

The CLT results for the simple mean (Kneip et al., 2021) and aggregate MPI (Pham et al., 2023) estimated via DEA are useful recent advancements of the statistical theory for efficiency and productivity analyses. However, for relatively small sample sizes and large dimensions, the coverages of the estimated confidence intervals based on the CLT results for the simple mean and aggregate MPI are below the nominal coverage. This under-covering phenomenon in relatively small sample sizes was also observed for the simple mean and aggregate efficiency.

Some of the improvements were made through Simar and Zelenyuk (2020), Nguyen et al. (2022), Simar et al. (2023a), and Simar et al. (2023b) for the simple mean and aggregate efficiency. However, whether these methods are effective to improve the finite sample performance of CLT results for the simple mean and aggregate MPI remains unknown. We fill this gap in the literature by thoroughly examining the performance of these two methods in Nguyen et al. (2022) and Simar et al. (2023a) for the simple mean and the aggregate MPI through extensive simulations and we also use one empirical data set to illustrate their differences.

In our extensive Monte-Carlo experiments, we find that the method adapted from Simar et al. (2023a) to the MPI context could provide a better performance for the simple mean and the aggregate MPI for relatively small samples (e.g., up to around 50, perhaps 100) and after that the original methods from Kneip et al. (2021) and Pham et al. (2023) are recommended.

Acknowledgement

Valentin Zelenyuk acknowledges the support from the Australian Research Council (FT170100401) and The University of Queensland. Shirong Zhao acknowledges the Liaoning Provincial Department of Education Research Fund (LJKMZ20221602). Feedback from Léopold Simar, Zhichao Wang and Evelyn Smart is appreciated. These individuals and organizations are not responsible for the views expressed here. These authors contributed equally: Valentin Zelenyuk, Shirong Zhao.

Declaration of Competing interest

The authors declare no known competing financial interests or personal relationships that could have appeared to influence this paper's results.

References

Badunenko, O., D. J. Henderson, and V. Zelenyuk (2008), Technological change and transition: relative contributions to worldwide growth during the 1990s, *Oxford Bulletin of Economics and Statistics* 70, 461–492.

— (2018), The productivity of nations, in E. Grifell-Tatjé, C. K. Lovell, and R. C. Sickles, eds., *The Oxford Handbook of Productivity Analysis*, Oxford University Press.

Casu, B., A. Ferrari, and T. Zhao (2013), Regulatory Reform and Productivity Change in Indian Banking, *The Review of Economics and Statistics* 95, 1066–1077.

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.

Färe, R., S. Grosskopf, M. Norris, and Z. Zhang (1994), 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.

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015), The next generation of the Penn World Table, *American Economic Review* 105, 3150–3182.

Kevork, I. S., J. Pange, P. Tzeremes, and N. G. Tzeremes (2017), Estimating Malmquist productivity indexes using probabilistic directional distances: An application to the European banking sector, *European Journal of Operational Research* 261, 1125–1140.

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.

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

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.

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

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 (2023), Statistical inference for aggregation of Malmquist productivity indices, *Operations Research* Forthcoming.

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 (2022), Another look at productivity growth in industrialized countries, Technical report, Université catholique de Louvain, Institute of Statistics, Biostatistics and Actuarial Sciences. LIDAM Discussion Papers ISBA.

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

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

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

— (2023b), Inference for aggregate efficiency: Theory and guidelines for practitioners. Centre for Efficiency and Productivity Analysis, Working Paper Series No. WP03/2023, School of Economics, University of Queensland St. Lucia, Qld. 4072, Australia.

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

Further Improvements of Finite Sample Approximation of Central Limit Theorems for Weighted and Unweighted Malmquist Productivity Indices (Supplementary Appendix)

Valentin Zelenyuk* Shirong Zhao†

May 2023

This is a draft not for circulation or citation without our permission

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

†Zhao: School of Finance, Dongbei University of Finance and Economics, Dalian, Liaoning 116025, China; email: shironz@dufe.edu.cn.

Contents

List of Tables

List of Figures

Appendix A Additional Simulation Results for $\delta = 0.04$

In this appendix, we present additional simulation results for $\delta = 0.04$, which are not shown in the paper.

Table A.1: Coverages of Estimated Confidence Intervals for the Simple Mean Malmquist Productivity Indices

			0.90					0.95					0.99				
p	q	n	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
1	1	20	0.894	0.894	0.894	0.879	0.879	0.949	0.949	0.949	0.937	0.937	0.991	0.991	0.991	0.984	0.984
1	1	50	0.903	0.903	0.903	0.903	0.903	0.953	0.953	0.953	0.951	0.951	0.990	0.990	0.990	0.990	0.990
1	1	100	0.888	0.888	0.888	0.888	0.888	0.946	0.946	0.946	0.944	0.944	0.985	0.985	0.985	0.984	0.984
1	1	200	0.892	0.892	0.892	0.891	0.891	0.943	0.943	0.943	0.943	0.943	0.985	0.985	0.985	0.985	0.985
1	1	300	0.900	0.900	0.900	0.899	0.899	0.939	0.939	0.939	0.940	0.940	0.988	0.988	0.988	0.987	0.987
1	1	500	0.896	0.896	0.896	0.896	0.896	0.943	0.943	0.943	0.943	0.943	0.990	0.990	0.990	0.989	0.989
1	1	1000	0.894	0.894	0.894	0.894	0.894	0.945	0.945	0.945	0.946	0.946	0.992	0.992	0.992	0.992	0.992
2	1	20	0.841	0.822	0.862	0.811	0.846	0.909	0.897	0.929	0.875	0.907	0.977	0.972	0.985	0.957	0.972
2	1	50	0.880	0.873	0.887	0.841	0.866	0.938	0.933	0.948	0.916	0.940	0.985	0.980	0.986	0.981	0.986
2	1	100	0.898	0.897	0.906	0.879	0.890	0.952	0.947	0.955	0.943	0.950	0.990	0.990	0.993	0.989	0.993
2	1	200	0.892	0.877	0.889	0.873	0.881	0.948	0.942	0.946	0.939	0.951	0.991	0.990	0.992	0.988	0.991
2	1	300	0.891	0.890	0.898	0.890	0.892	0.949	0.950	0.957	0.941	0.946	0.990	0.991	0.991	0.990	0.992
2	1	500	0.893	0.889	0.892	0.884	0.890	0.948	0.943	0.945	0.940	0.944	0.985	0.985	0.985	0.985	0.988
2	1	1000	0.881	0.880	0.882	0.882	0.884	0.947	0.946	0.948	0.945	0.949	0.993	0.990	0.990	0.990	0.991
3	1	20	0.810	0.812	0.898	0.841	0.891	0.896	0.881	0.949	0.905	0.947	0.961	0.964	0.995	0.978	0.988
3	1	50	0.888	0.863	0.911	0.843	0.886	0.928	0.920	0.963	0.911	0.940	0.982	0.982	0.993	0.968	0.987
3	1	100	0.866	0.870	0.909	0.860	0.894	0.936	0.939	0.963	0.913	0.948	0.986	0.986	0.992	0.980	0.989
3	1	200	0.899	0.907	0.926	0.887	0.910	0.955	0.950	0.967	0.937	0.952	0.991	0.992	0.994	0.990	0.993
3	1	300	0.885	0.888	0.908	0.882	0.901	0.936	0.945	0.956	0.930	0.945	0.991	0.987	0.990	0.983	0.988
3	1	500	0.890	0.908	0.922	0.902	0.912	0.949	0.958	0.965	0.954	0.962	0.986	0.988	0.991	0.987	0.989
3	1	1000	0.905	0.899	0.906	0.890	0.897	0.952	0.946	0.951	0.943	0.947	0.986	0.990	0.992	0.990	0.992

Table A.1: Coverages of Estimated Confidence Intervals for the Simple Mean Malmquist Productivity Indices (continued)

			0.90					0.95					0.99				
p	q	n	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
4	1	20	0.808	0.774	0.902	0.842	0.905	0.876	0.858	0.959	0.910	0.950	0.962	0.956	0.993	0.966	0.984
4	1	50	0.840	0.868	0.935	0.870	0.923	0.899	0.931	0.968	0.924	0.967	0.964	0.975	0.993	0.980	0.992
4	1	100	0.871	0.881	0.938	0.872	0.918	0.926	0.935	0.982	0.921	0.966	0.972	0.990	0.998	0.981	0.996
4	1	200	0.881	0.918	0.947	0.892	0.943	0.942	0.959	0.979	0.950	0.973	0.987	0.985	0.994	0.984	0.995
4	1	300	0.890	0.898	0.928	0.869	0.913	0.942	0.936	0.960	0.928	0.960	0.986	0.986	0.991	0.981	0.989
4	1	500	0.893	0.916	0.937	0.907	0.934	0.954	0.956	0.972	0.957	0.976	0.991	0.989	0.994	0.990	0.994
4	1	1000	0.888	0.906	0.929	0.898	0.919	0.945	0.961	0.967	0.953	0.964	0.985	0.989	0.992	0.987	0.994
5	1	20	0.773	0.783	0.910	0.859	0.900	0.844	0.863	0.957	0.912	0.943	0.936	0.948	0.989	0.971	0.985
5	1	50	0.826	0.862	0.946	0.867	0.924	0.887	0.920	0.975	0.921	0.965	0.958	0.976	0.995	0.978	0.989
5	1	100	0.852	0.874	0.946	0.893	0.944	0.911	0.934	0.980	0.942	0.971	0.979	0.978	0.994	0.982	0.993
5	1	200	0.866	0.889	0.945	0.887	0.944	0.918	0.941	0.981	0.944	0.975	0.978	0.990	1.000	0.989	0.998
5	1	300	0.878	0.890	0.945	0.903	0.944	0.930	0.949	0.974	0.947	0.969	0.982	0.988	0.994	0.985	0.996
5	1	500	0.877	0.902	0.938	0.892	0.939	0.931	0.951	0.978	0.947	0.974	0.983	0.990	0.998	0.990	0.997
5	1	1000	0.887	0.895	0.929	0.889	0.925	0.941	0.946	0.967	0.942	0.967	0.987	0.992	0.994	0.988	0.997
7	1	20	0.770	0.761	0.922	0.864	0.915	0.842	0.838	0.959	0.932	0.967	0.930	0.938	0.986	0.988	0.995
7	1	50	0.803	0.859	0.956	0.868	0.921	0.874	0.913	0.981	0.927	0.969	0.943	0.970	0.998	0.981	0.996
7	1	100	0.792	0.894	0.968	0.871	0.945	0.865	0.938	0.989	0.932	0.975	0.949	0.981	0.996	0.987	0.996
7	1	200	0.808	0.891	0.971	0.897	0.959	0.875	0.943	0.986	0.945	0.982	0.955	0.980	0.998	0.987	0.999
7	1	300	0.812	0.896	0.968	0.896	0.962	0.880	0.944	0.987	0.945	0.989	0.963	0.983	0.998	0.994	1.000
7	1	500	0.816	0.906	0.960	0.907	0.963	0.882	0.946	0.980	0.952	0.988	0.967	0.980	1.000	0.992	0.999
7	1	1000	0.807	0.883	0.942	0.890	0.950	0.883	0.930	0.974	0.942	0.981	0.960	0.981	0.997	0.987	0.997

Table A.2: Coverages of Estimated Confidence Intervals for the Aggregate Malmquist Productivity Indices

					0.90					0.95					0.99		
p	q	n	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
1	1	20	0.890	0.894	0.889	0.879	0.879	0.950	0.947	0.948	0.943	0.944	0.989	0.991	0.989	0.983	0.982
1	1	50	0.902	0.899	0.897	0.899	0.899	0.942	0.942	0.942	0.943	0.946	0.982	0.981	0.982	0.980	0.982
1	1	100	0.884	0.882	0.882	0.883	0.883	0.942	0.942	0.941	0.941	0.941	0.983	0.981	0.982	0.981	0.982
1	1	200	0.901	0.903	0.902	0.903	0.902	0.944	0.943	0.943	0.943	0.943	0.985	0.985	0.985	0.985	0.985
1	1	300	0.902	0.903	0.903	0.903	0.902	0.955	0.955	0.955	0.955	0.955	0.990	0.990	0.990	0.990	0.990
1	1	500	0.893	0.891	0.891	0.891	0.891	0.935	0.936	0.937	0.937	0.937	0.983	0.983	0.983	0.983	0.983
1	1	1000	0.875	0.873	0.873	0.873	0.873	0.945	0.944	0.944	0.944	0.944	0.985	0.984	0.984	0.984	0.984
2	1	20	0.860	0.822	0.869	0.801	0.837	0.918	0.884	0.919	0.876	0.895	0.969	0.960	0.973	0.944	0.960
2	1	50	0.879	0.866	0.882	0.825	0.853	0.932	0.923	0.941	0.892	0.914	0.984	0.986	0.988	0.963	0.973
2	1	100	0.902	0.901	0.905	0.867	0.877	0.947	0.945	0.948	0.919	0.922	0.990	0.988	0.991	0.975	0.978
2	1	200	0.891	0.880	0.890	0.853	0.861	0.947	0.946	0.949	0.927	0.934	0.990	0.991	0.993	0.984	0.986
2	1	300	0.903	0.896	0.899	0.886	0.888	0.958	0.955	0.960	0.937	0.942	0.990	0.993	0.993	0.987	0.988
2	1	500	0.901	0.896	0.897	0.881	0.885	0.941	0.945	0.945	0.942	0.944	0.990	0.988	0.989	0.984	0.986
2	1	1000	0.893	0.899	0.902	0.885	0.887	0.952	0.951	0.952	0.948	0.950	0.988	0.989	0.989	0.986	0.987
3	1	20	0.823	0.831	0.909	0.845	0.904	0.894	0.898	0.964	0.919	0.942	0.966	0.972	0.994	0.970	0.985
3	1	50	0.872	0.872	0.923	0.838	0.886	0.928	0.936	0.964	0.899	0.938	0.987	0.984	0.991	0.961	0.987
3	1	100	0.883	0.884	0.906	0.833	0.878	0.937	0.942	0.960	0.911	0.932	0.989	0.983	0.992	0.968	0.979
3	1	200	0.891	0.913	0.930	0.847	0.868	0.946	0.960	0.967	0.914	0.933	0.991	0.995	0.997	0.982	0.987
3	1	300	0.893	0.899	0.908	0.860	0.868	0.940	0.952	0.963	0.925	0.932	0.989	0.993	0.994	0.983	0.988
3	1	500	0.897	0.888	0.898	0.886	0.891	0.939	0.954	0.967	0.930	0.940	0.980	0.993	0.995	0.982	0.983
3	1	1000	0.907	0.894	0.896	0.883	0.889	0.953	0.950	0.956	0.947	0.950	0.991	0.992	0.992	0.988	0.990

Table A.2: Coverages of Estimated Confidence Intervals for the Aggregate Malmquist Productivity Indices (continued)

			0.90					0.95					0.99				
p	q	n	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
4	1	20	0.797	0.779	0.920	0.836	0.890	0.869	0.862	0.957	0.889	0.937	0.948	0.947	0.992	0.964	0.982
4	1	50	0.832	0.868	0.941	0.878	0.927	0.905	0.918	0.972	0.934	0.972	0.971	0.974	0.996	0.980	0.996
4	1	100	0.860	0.875	0.938	0.878	0.926	0.917	0.929	0.974	0.929	0.963	0.969	0.982	0.997	0.977	0.997
4	1	200	0.892	0.909	0.948	0.890	0.939	0.950	0.952	0.983	0.942	0.973	0.988	0.989	0.993	0.985	0.995
4	1	300	0.881	0.899	0.938	0.892	0.933	0.946	0.953	0.975	0.940	0.973	0.992	0.988	0.998	0.988	0.995
4	1	500	0.906	0.900	0.925	0.887	0.927	0.953	0.950	0.965	0.944	0.966	0.986	0.990	0.995	0.988	0.994
4	1	1000	0.888	0.915	0.926	0.911	0.929	0.944	0.959	0.965	0.951	0.961	0.991	0.991	0.998	0.993	0.997
5	1	20	0.756	0.779	0.912	0.861	0.914	0.840	0.847	0.950	0.924	0.947	0.927	0.934	0.989	0.975	0.983
5	1	50	0.823	0.854	0.955	0.862	0.931	0.881	0.914	0.982	0.933	0.969	0.954	0.971	0.998	0.978	0.989
5	1	100	0.836	0.868	0.960	0.892	0.949	0.907	0.928	0.988	0.948	0.976	0.974	0.985	0.998	0.983	0.997
5	1	200	0.848	0.898	0.970	0.882	0.945	0.911	0.950	0.992	0.943	0.980	0.973	0.995	0.999	0.989	0.999
5	1	300	0.871	0.888	0.955	0.896	0.950	0.930	0.948	0.982	0.944	0.978	0.980	0.986	0.997	0.986	0.996
5	1	500	0.868	0.885	0.946	0.868	0.943	0.938	0.949	0.979	0.940	0.974	0.977	0.991	1.000	0.986	1.000
5	1	1000	0.892	0.901	0.948	0.879	0.937	0.948	0.957	0.978	0.943	0.973	0.983	0.989	0.997	0.987	0.999
7	1	20	0.763	0.741	0.920	0.867	0.905	0.834	0.825	0.952	0.924	0.957	0.923	0.923	0.986	0.985	0.994
7	1	50	0.799	0.844	0.960	0.863	0.919	0.870	0.907	0.985	0.923	0.956	0.946	0.968	0.997	0.984	0.993
7	1	100	0.794	0.868	0.976	0.879	0.941	0.869	0.925	0.994	0.933	0.967	0.957	0.978	1.000	0.984	0.996
7	1	200	0.830	0.887	0.987	0.899	0.957	0.900	0.944	0.996	0.945	0.986	0.960	0.989	1.000	0.990	0.998
7	1	300	0.844	0.886	0.982	0.901	0.958	0.915	0.942	0.992	0.951	0.986	0.972	0.989	0.998	0.990	0.999
7	1	500	0.852	0.912	0.971	0.911	0.969	0.921	0.953	0.991	0.958	0.992	0.981	0.989	0.999	0.992	1.000
7	1	1000	0.846	0.885	0.960	0.901	0.959	0.906	0.942	0.983	0.956	0.985	0.966	0.986	0.999	0.989	0.998

Appendix B Simulation Results for $\delta = 0.02$

We present the simulation results for the case when $\delta = 0.02$ here.

Figure B.1: Coverages of Estimated Confidence Intervals for the Simple Mean Malmquist Productivity Indices for the 90% Nominal Coverage, with $\delta = 0.02$

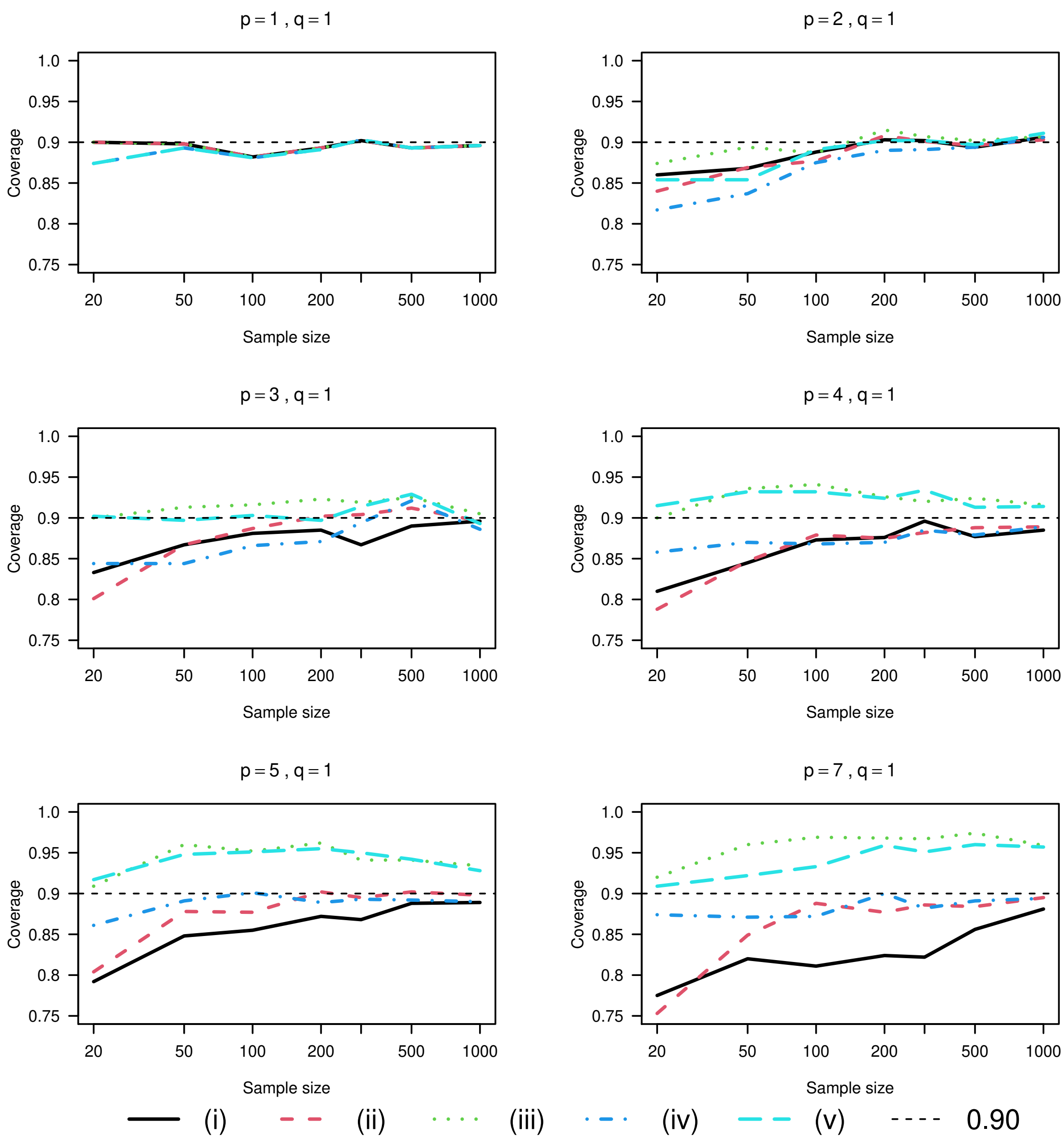


Figure B.2: Coverages of Estimated Confidence Intervals for the Simple Mean Malmquist Productivity Indices for the 95% Nominal Coverage, with $\delta = 0.02$

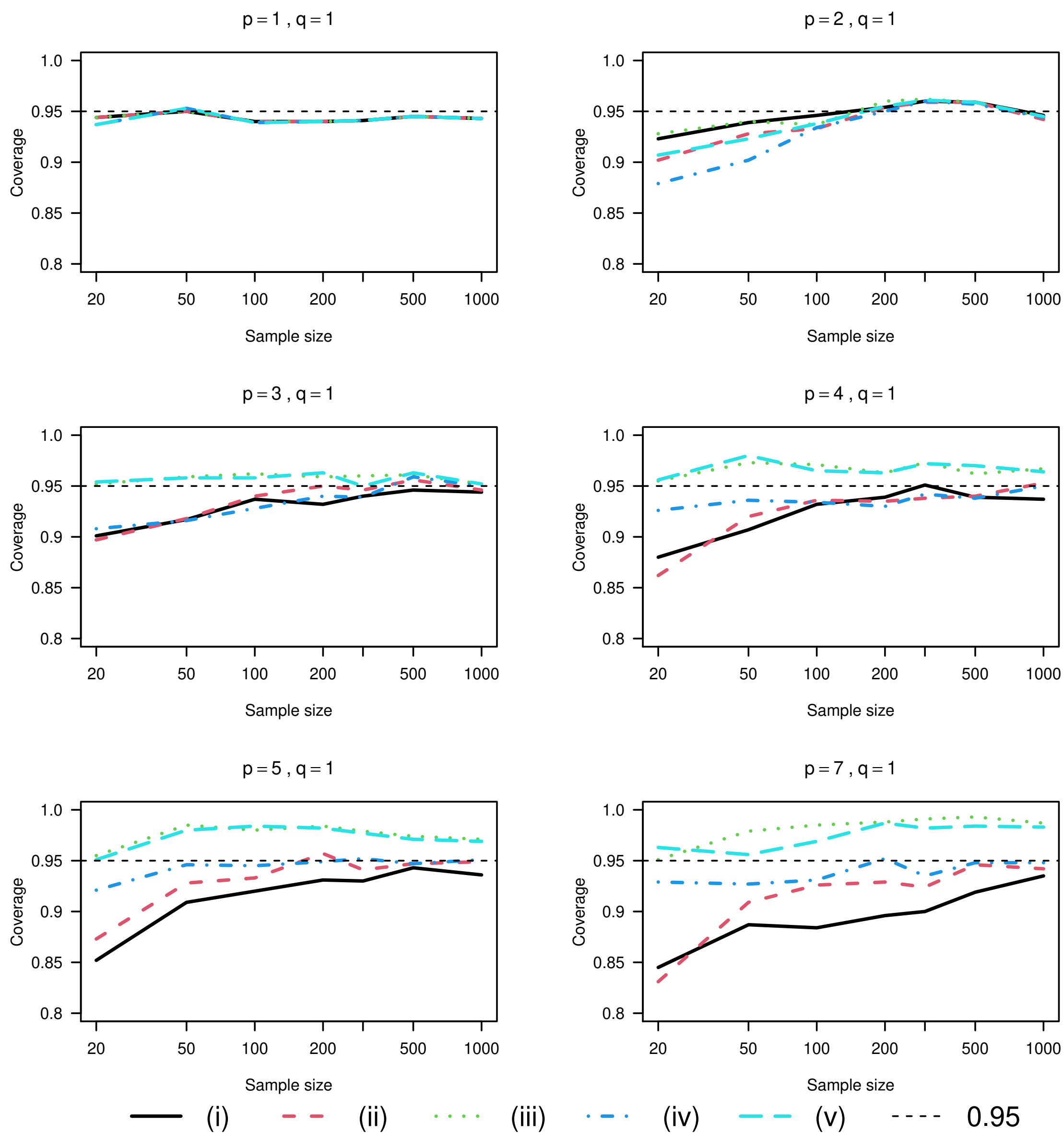


Figure B.3: Coverages of Estimated Confidence Intervals for the Aggregate Malmquist Productivity Indices for the 90% Nominal Coverage, with $\delta = 0.02$

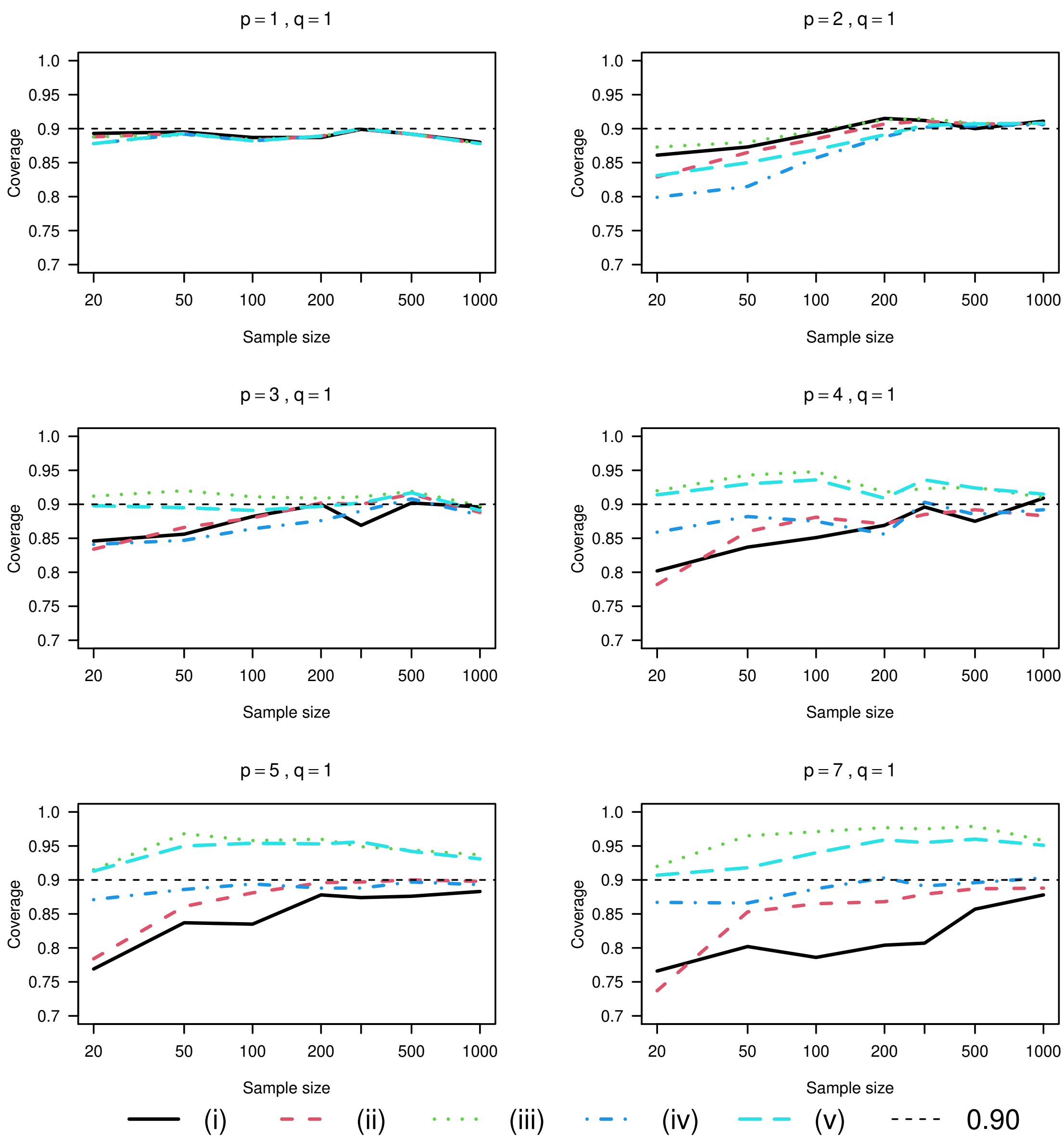


Figure B.4: Coverages of Estimated Confidence Intervals for the Aggregate Malmquist Productivity Indices for the 95% Nominal Coverage, with $\delta = 0.02$

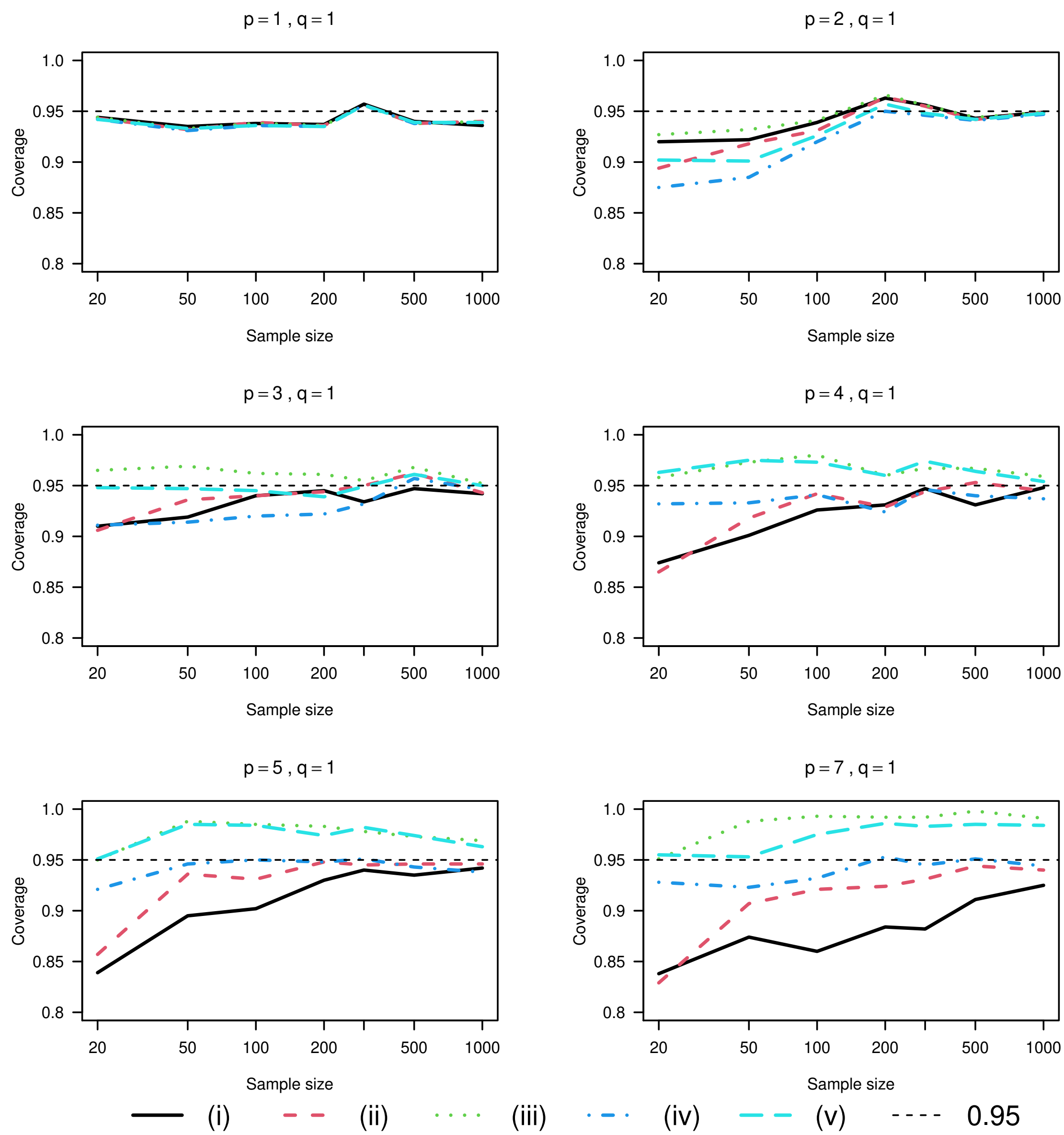


Table B.1: Coverages of Estimated Confidence Intervals for the Simple Mean Malmquist Productivity Indices, with $\delta = 0.02$

			0.90					0.95					0.99				
p	q	n	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
1	1	20	0.900	0.900	0.900	0.874	0.874	0.944	0.944	0.944	0.937	0.937	0.985	0.985	0.985	0.978	0.978
1	1	50	0.898	0.898	0.898	0.893	0.893	0.950	0.950	0.950	0.953	0.953	0.991	0.991	0.991	0.991	0.991
1	1	100	0.882	0.882	0.882	0.881	0.881	0.940	0.940	0.940	0.939	0.939	0.985	0.985	0.985	0.986	0.986
1	1	200	0.893	0.893	0.893	0.891	0.891	0.940	0.940	0.940	0.940	0.940	0.984	0.984	0.984	0.984	0.984
1	1	300	0.902	0.902	0.902	0.903	0.903	0.941	0.941	0.941	0.941	0.941	0.988	0.988	0.988	0.988	0.988
1	1	500	0.893	0.893	0.893	0.893	0.893	0.945	0.945	0.945	0.945	0.945	0.990	0.990	0.990	0.990	0.990
1	1	1000	0.896	0.896	0.896	0.896	0.896	0.943	0.943	0.943	0.943	0.943	0.988	0.988	0.988	0.988	0.988
2	1	20	0.860	0.840	0.874	0.817	0.854	0.923	0.902	0.928	0.879	0.907	0.976	0.964	0.985	0.949	0.972
2	1	50	0.868	0.869	0.894	0.837	0.854	0.939	0.928	0.940	0.902	0.923	0.983	0.981	0.987	0.967	0.976
2	1	100	0.888	0.877	0.887	0.875	0.891	0.946	0.933	0.937	0.934	0.938	0.992	0.989	0.991	0.982	0.985
2	1	200	0.903	0.908	0.915	0.890	0.902	0.954	0.954	0.960	0.950	0.955	0.991	0.990	0.990	0.989	0.990
2	1	300	0.902	0.901	0.907	0.891	0.902	0.960	0.959	0.962	0.960	0.961	0.993	0.991	0.993	0.991	0.993
2	1	500	0.894	0.896	0.902	0.894	0.897	0.959	0.959	0.959	0.957	0.959	0.990	0.990	0.990	0.989	0.989
2	1	1000	0.906	0.903	0.907	0.906	0.911	0.946	0.942	0.944	0.944	0.945	0.989	0.989	0.989	0.989	0.989
3	1	20	0.833	0.801	0.900	0.844	0.902	0.901	0.897	0.952	0.908	0.954	0.970	0.968	0.991	0.976	0.988
3	1	50	0.867	0.867	0.913	0.844	0.897	0.917	0.918	0.959	0.916	0.958	0.975	0.983	0.995	0.980	0.989
3	1	100	0.881	0.887	0.916	0.866	0.903	0.937	0.940	0.962	0.928	0.958	0.988	0.990	0.995	0.989	0.994
3	1	200	0.885	0.902	0.923	0.871	0.897	0.932	0.950	0.959	0.940	0.963	0.991	0.992	0.996	0.991	0.996
3	1	300	0.867	0.904	0.919	0.894	0.914	0.940	0.946	0.960	0.939	0.950	0.981	0.987	0.990	0.992	0.993
3	1	500	0.890	0.912	0.925	0.921	0.929	0.946	0.956	0.961	0.959	0.963	0.989	0.991	0.993	0.991	0.992
3	1	1000	0.896	0.896	0.905	0.886	0.893	0.944	0.946	0.952	0.950	0.952	0.985	0.990	0.991	0.994	0.995

Table B.1: Coverages of Estimated Confidence Intervals for the Simple Mean Malmquist Productivity Indices, with $\delta = 0.02$ (continued)

					0.90					0.95					0.99		
p	q	n	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
4	1	20	0.810	0.788	0.900	0.858	0.915	0.880	0.862	0.955	0.926	0.956	0.964	0.952	0.987	0.976	0.989
4	1	50	0.845	0.847	0.936	0.870	0.932	0.907	0.920	0.973	0.936	0.980	0.970	0.983	0.997	0.985	0.996
4	1	100	0.873	0.879	0.941	0.868	0.932	0.932	0.936	0.971	0.934	0.965	0.978	0.979	0.995	0.978	0.998
4	1	200	0.876	0.875	0.926	0.870	0.924	0.939	0.935	0.963	0.930	0.963	0.985	0.980	0.993	0.985	0.991
4	1	300	0.896	0.882	0.920	0.885	0.934	0.951	0.938	0.973	0.942	0.972	0.989	0.994	0.999	0.986	0.997
4	1	500	0.877	0.888	0.924	0.879	0.913	0.939	0.940	0.962	0.938	0.970	0.984	0.988	0.995	0.989	0.995
4	1	1000	0.885	0.889	0.916	0.889	0.914	0.937	0.953	0.967	0.950	0.964	0.988	0.993	0.994	0.990	0.993
5	1	20	0.792	0.804	0.909	0.861	0.917	0.852	0.873	0.955	0.921	0.951	0.939	0.946	0.987	0.976	0.986
5	1	50	0.848	0.878	0.960	0.891	0.948	0.909	0.928	0.985	0.946	0.980	0.972	0.982	0.996	0.986	0.996
5	1	100	0.855	0.877	0.952	0.901	0.951	0.920	0.933	0.980	0.945	0.984	0.977	0.982	0.998	0.994	0.999
5	1	200	0.872	0.902	0.962	0.889	0.955	0.931	0.957	0.984	0.949	0.982	0.980	0.988	0.997	0.990	0.996
5	1	300	0.868	0.895	0.941	0.893	0.950	0.930	0.941	0.979	0.952	0.977	0.981	0.991	0.999	0.990	0.997
5	1	500	0.888	0.902	0.941	0.892	0.942	0.943	0.947	0.974	0.947	0.971	0.984	0.986	0.994	0.986	0.994
5	1	1000	0.889	0.898	0.934	0.890	0.928	0.936	0.949	0.971	0.951	0.969	0.984	0.985	0.994	0.987	0.996
7	1	20	0.775	0.753	0.920	0.874	0.909	0.845	0.831	0.951	0.929	0.963	0.933	0.932	0.987	0.987	0.992
7	1	50	0.820	0.849	0.960	0.871	0.922	0.887	0.909	0.979	0.927	0.956	0.945	0.970	0.999	0.980	0.995
7	1	100	0.811	0.888	0.969	0.872	0.933	0.884	0.926	0.985	0.931	0.969	0.951	0.977	0.995	0.985	0.997
7	1	200	0.824	0.877	0.968	0.900	0.959	0.896	0.929	0.988	0.952	0.987	0.969	0.983	0.999	0.990	0.998
7	1	300	0.822	0.886	0.967	0.882	0.951	0.900	0.924	0.991	0.935	0.982	0.973	0.986	1.000	0.990	1.000
7	1	500	0.856	0.884	0.974	0.891	0.960	0.919	0.946	0.993	0.948	0.984	0.979	0.989	1.000	0.988	0.998
7	1	1000	0.881	0.895	0.959	0.894	0.957	0.935	0.942	0.987	0.948	0.983	0.987	0.989	0.998	0.990	0.999

Table B.2: Coverages of Estimated Confidence Intervals for the Aggregate Malmquist Productivity Indices, with $\delta = 0.02$

			0.90					0.95					0.99				
p	q	n	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
1	1	20	0.893	0.888	0.888	0.878	0.878	0.944	0.943	0.944	0.942	0.942	0.986	0.989	0.988	0.984	0.981
1	1	50	0.895	0.893	0.891	0.892	0.893	0.935	0.932	0.933	0.931	0.933	0.989	0.987	0.988	0.986	0.987
1	1	100	0.887	0.883	0.884	0.882	0.882	0.938	0.939	0.938	0.936	0.936	0.979	0.979	0.980	0.977	0.979
1	1	200	0.887	0.889	0.889	0.889	0.889	0.937	0.936	0.936	0.935	0.935	0.984	0.984	0.984	0.984	0.984
1	1	300	0.899	0.899	0.899	0.900	0.900	0.957	0.956	0.956	0.956	0.956	0.988	0.988	0.988	0.988	0.988
1	1	500	0.892	0.892	0.892	0.892	0.892	0.940	0.938	0.938	0.938	0.939	0.983	0.983	0.983	0.983	0.983
1	1	1000	0.880	0.878	0.878	0.878	0.878	0.936	0.940	0.940	0.939	0.939	0.986	0.986	0.986	0.986	0.986
2	1	20	0.861	0.829	0.873	0.799	0.831	0.920	0.894	0.927	0.875	0.902	0.977	0.963	0.978	0.943	0.965
2	1	50	0.873	0.865	0.880	0.815	0.850	0.922	0.918	0.932	0.885	0.901	0.981	0.982	0.988	0.959	0.970
2	1	100	0.893	0.885	0.897	0.857	0.869	0.939	0.931	0.941	0.920	0.926	0.990	0.988	0.990	0.983	0.984
2	1	200	0.915	0.907	0.913	0.888	0.891	0.963	0.964	0.966	0.950	0.957	0.991	0.991	0.992	0.988	0.988
2	1	300	0.912	0.911	0.915	0.902	0.905	0.956	0.955	0.956	0.946	0.948	0.992	0.992	0.992	0.993	0.993
2	1	500	0.900	0.907	0.908	0.904	0.907	0.943	0.942	0.944	0.941	0.942	0.984	0.983	0.985	0.982	0.984
2	1	1000	0.911	0.907	0.907	0.906	0.908	0.949	0.949	0.949	0.947	0.948	0.988	0.988	0.988	0.988	0.988
3	1	20	0.846	0.834	0.912	0.841	0.898	0.910	0.906	0.965	0.911	0.948	0.971	0.970	0.995	0.975	0.991
3	1	50	0.856	0.866	0.920	0.847	0.895	0.919	0.936	0.969	0.914	0.947	0.977	0.985	0.994	0.978	0.993
3	1	100	0.882	0.880	0.911	0.864	0.891	0.940	0.940	0.962	0.920	0.945	0.986	0.984	0.993	0.980	0.990
3	1	200	0.900	0.902	0.909	0.876	0.897	0.945	0.944	0.961	0.922	0.939	0.984	0.990	0.996	0.984	0.992
3	1	300	0.869	0.900	0.911	0.890	0.902	0.934	0.950	0.955	0.932	0.949	0.992	0.990	0.992	0.980	0.987
3	1	500	0.902	0.915	0.919	0.908	0.917	0.947	0.962	0.968	0.957	0.961	0.993	0.992	0.992	0.991	0.992
3	1	1000	0.896	0.888	0.898	0.884	0.891	0.942	0.943	0.952	0.946	0.950	0.985	0.984	0.986	0.984	0.985

Table B.2: Coverages of Estimated Confidence Intervals for the Aggregate Malmquist Productivity Indices, with $\delta = 0.02$ (continued)

			0.90					0.95					0.99				
p	q	n	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
4	1	20	0.802	0.782	0.920	0.859	0.914	0.874	0.865	0.958	0.932	0.963	0.958	0.943	0.994	0.979	0.990
4	1	50	0.837	0.860	0.943	0.882	0.930	0.901	0.918	0.973	0.933	0.975	0.972	0.976	0.996	0.989	0.999
4	1	100	0.851	0.881	0.948	0.875	0.936	0.926	0.942	0.980	0.941	0.973	0.983	0.987	0.999	0.984	0.997
4	1	200	0.869	0.871	0.918	0.856	0.909	0.931	0.929	0.960	0.924	0.960	0.988	0.986	0.993	0.981	0.991
4	1	300	0.896	0.885	0.923	0.903	0.936	0.947	0.944	0.967	0.946	0.974	0.985	0.992	0.998	0.982	0.993
4	1	500	0.875	0.892	0.925	0.885	0.924	0.931	0.953	0.967	0.940	0.964	0.977	0.990	0.995	0.985	0.997
4	1	1000	0.909	0.883	0.910	0.892	0.915	0.948	0.945	0.959	0.937	0.954	0.989	0.989	0.992	0.991	0.993
5	1	20	0.769	0.784	0.915	0.871	0.913	0.839	0.857	0.950	0.921	0.951	0.929	0.938	0.988	0.973	0.983
5	1	50	0.837	0.861	0.968	0.886	0.950	0.895	0.936	0.988	0.946	0.985	0.971	0.979	0.999	0.990	0.996
5	1	100	0.835	0.881	0.958	0.894	0.954	0.902	0.931	0.985	0.950	0.984	0.971	0.982	0.998	0.989	1.000
5	1	200	0.878	0.896	0.960	0.888	0.953	0.930	0.948	0.983	0.948	0.974	0.986	0.985	0.998	0.990	0.998
5	1	300	0.874	0.897	0.949	0.888	0.956	0.940	0.945	0.978	0.951	0.982	0.984	0.989	0.999	0.990	0.997
5	1	500	0.876	0.900	0.944	0.897	0.942	0.935	0.946	0.973	0.943	0.974	0.980	0.989	0.998	0.982	0.998
5	1	1000	0.883	0.898	0.937	0.893	0.931	0.942	0.946	0.969	0.938	0.963	0.985	0.989	0.994	0.984	0.994
7	1	20	0.766	0.737	0.920	0.867	0.907	0.838	0.829	0.951	0.928	0.955	0.937	0.927	0.988	0.987	0.998
7	1	50	0.802	0.853	0.965	0.866	0.918	0.874	0.907	0.988	0.923	0.953	0.943	0.971	0.998	0.977	0.993
7	1	100	0.786	0.865	0.971	0.887	0.940	0.860	0.921	0.993	0.932	0.975	0.952	0.970	0.999	0.982	0.995
7	1	200	0.804	0.868	0.977	0.903	0.959	0.884	0.924	0.992	0.953	0.986	0.965	0.985	1.000	0.992	0.999
7	1	300	0.807	0.879	0.975	0.891	0.955	0.882	0.931	0.992	0.945	0.983	0.961	0.985	1.000	0.992	0.999
7	1	500	0.857	0.887	0.979	0.896	0.960	0.911	0.944	0.998	0.951	0.985	0.973	0.993	1.000	0.993	0.998
7	1	1000	0.878	0.888	0.958	0.903	0.951	0.925	0.940	0.991	0.944	0.984	0.979	0.991	0.999	0.991	0.999

Appendix C Simulation Results for $\delta = 0.10$

We present the simulation results for the case when $\delta = 0.10$ here. Note that, to ensure the VRS for the technology, we have changed the values of βs, which are presented in Table C.1.

Table C.1: The Values of β_j and w_j

p	1	2	3	4	5	7
β_1	0.5	0.3	0.1	0.1	0.05	0.025
β_2		0.4	0.2	0.1	0.05	0.025
β_3			0.3	0.15	0.1	0.025
β_4				0.2	0.1	0.025
β_5					0.15	0.05
β_6						0.05
β_7						0.05

Figure C.1: Coverages of Estimated Confidence Intervals for the Simple Mean Malmquist Productivity Indices for the 90% Nominal Coverage, with $\delta = 0.10$

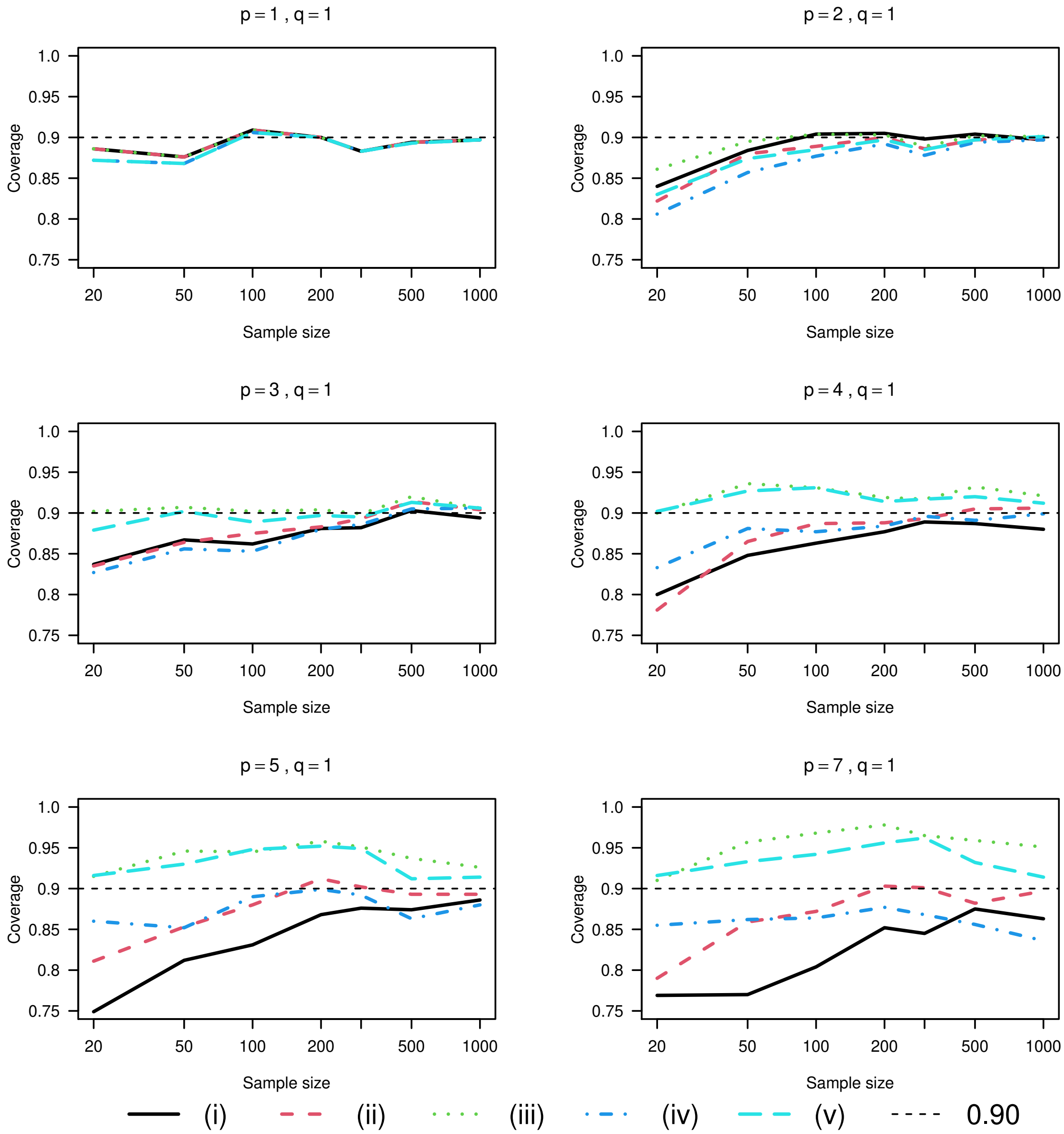


Figure C.2: Coverages of Estimated Confidence Intervals for the Simple Mean Malmquist Productivity Indices for the 95% Nominal Coverage, with $\delta = 0.10$

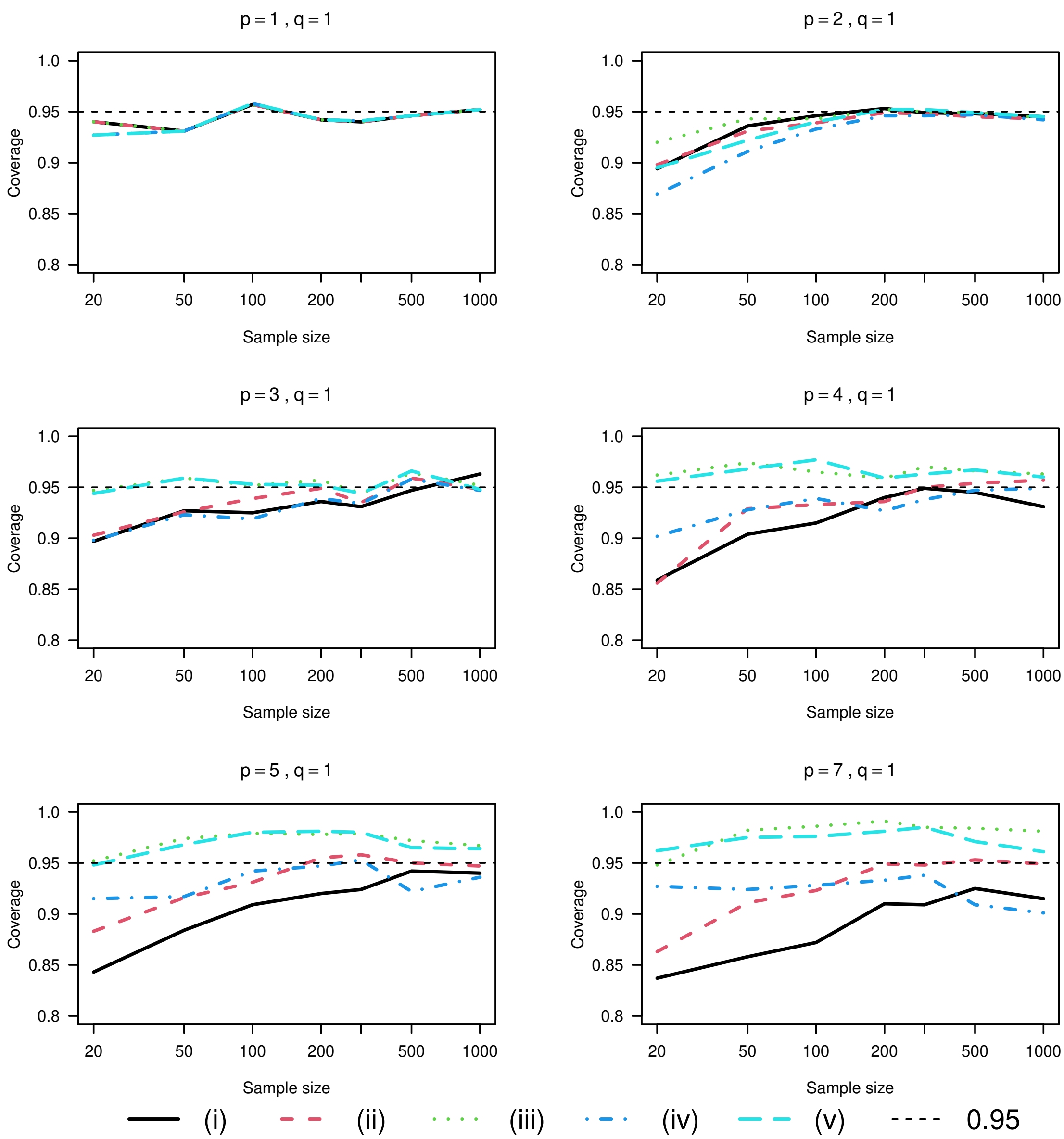


Figure C.3: Coverages of Estimated Confidence Intervals for the Aggregate Malmquist Productivity Indices for the 90% Nominal Coverage, with $\delta = 0.10$

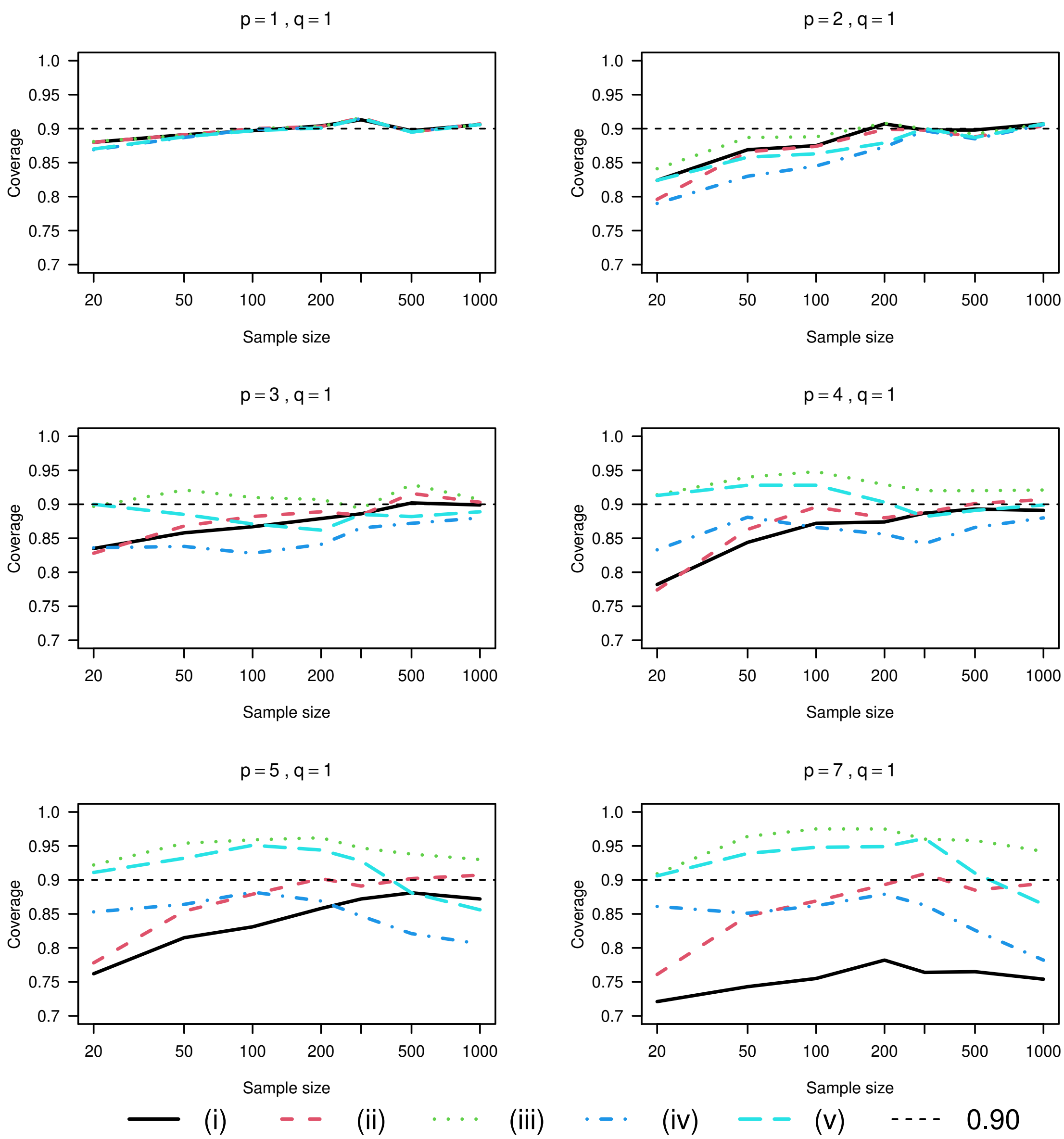


Figure C.4: Coverages of Estimated Confidence Intervals for the Aggregate Malmquist Productivity Indices for the 95% Nominal Coverage, with $\delta = 0.10$

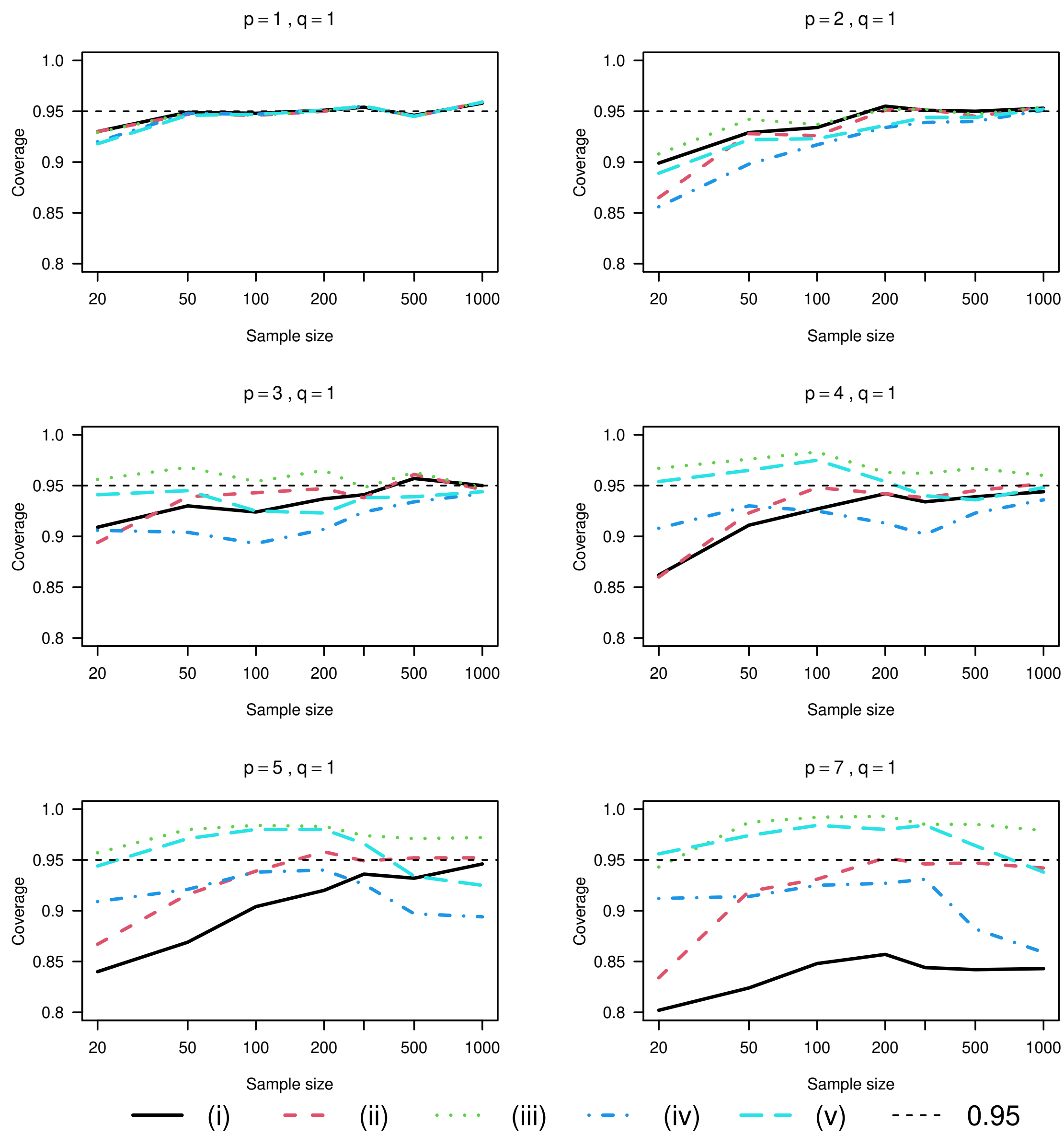


Table C.2: Coverages of Estimated Confidence Intervals for the Simple Mean Malmquist Productivity Indices, with $\delta = 0.10$

			0.90					0.95					0.99				
p	q	n	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
1	1	20	0.886	0.886	0.886	0.872	0.872	0.940	0.940	0.940	0.927	0.927	0.983	0.983	0.983	0.977	0.977
1	1	50	0.876	0.876	0.876	0.868	0.868	0.931	0.931	0.931	0.931	0.931	0.984	0.984	0.984	0.987	0.987
1	1	100	0.909	0.909	0.909	0.906	0.906	0.957	0.957	0.957	0.958	0.958	0.988	0.988	0.988	0.988	0.988
1	1	200	0.900	0.900	0.900	0.900	0.900	0.942	0.942	0.942	0.942	0.942	0.992	0.992	0.992	0.991	0.991
1	1	300	0.883	0.883	0.883	0.883	0.883	0.940	0.940	0.940	0.941	0.941	0.987	0.987	0.987	0.987	0.987
1	1	500	0.894	0.894	0.894	0.893	0.893	0.946	0.946	0.946	0.946	0.946	0.987	0.987	0.987	0.987	0.987
1	1	1000	0.897	0.897	0.897	0.897	0.897	0.952	0.952	0.952	0.952	0.952	0.989	0.989	0.989	0.989	0.989
2	1	20	0.840	0.822	0.861	0.806	0.830	0.894	0.898	0.920	0.869	0.895	0.961	0.959	0.970	0.940	0.961
2	1	50	0.884	0.880	0.895	0.857	0.874	0.936	0.931	0.943	0.911	0.922	0.984	0.986	0.987	0.976	0.979
2	1	100	0.904	0.889	0.904	0.877	0.885	0.946	0.939	0.943	0.933	0.940	0.985	0.984	0.987	0.980	0.982
2	1	200	0.905	0.900	0.903	0.892	0.897	0.953	0.949	0.952	0.946	0.952	0.990	0.992	0.994	0.991	0.993
2	1	300	0.898	0.886	0.889	0.878	0.885	0.949	0.948	0.950	0.946	0.952	0.984	0.985	0.985	0.985	0.986
2	1	500	0.904	0.897	0.902	0.894	0.897	0.948	0.945	0.949	0.947	0.949	0.986	0.985	0.985	0.983	0.984
2	1	1000	0.897	0.899	0.901	0.897	0.901	0.945	0.943	0.944	0.942	0.945	0.994	0.993	0.993	0.993	0.993
3	1	20	0.837	0.835	0.902	0.827	0.879	0.897	0.903	0.947	0.898	0.944	0.962	0.967	0.989	0.974	0.990
3	1	50	0.867	0.864	0.907	0.856	0.902	0.927	0.926	0.959	0.923	0.959	0.980	0.986	0.996	0.981	0.992
3	1	100	0.862	0.875	0.902	0.853	0.889	0.925	0.939	0.952	0.919	0.953	0.984	0.986	0.992	0.983	0.992
3	1	200	0.881	0.883	0.904	0.880	0.897	0.936	0.949	0.957	0.939	0.952	0.986	0.989	0.993	0.987	0.992
3	1	300	0.882	0.893	0.899	0.886	0.895	0.931	0.935	0.944	0.934	0.944	0.985	0.984	0.991	0.985	0.987
3	1	500	0.903	0.914	0.920	0.905	0.913	0.947	0.959	0.964	0.958	0.966	0.992	0.993	0.993	0.991	0.993
3	1	1000	0.894	0.904	0.906	0.906	0.906	0.963	0.948	0.952	0.947	0.948	0.992	0.990	0.990	0.989	0.989

Table C.2: Coverages of Estimated Confidence Intervals for the Simple Mean Malmquist Productivity Indices, with $\delta = 0.10$ (continued)

			0.90					0.95					0.99				
p	q	n	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
4	1	20	0.800	0.781	0.900	0.833	0.902	0.859	0.856	0.962	0.902	0.956	0.942	0.958	0.994	0.976	0.991
4	1	50	0.848	0.865	0.936	0.881	0.927	0.904	0.929	0.974	0.928	0.968	0.969	0.982	0.998	0.977	0.995
4	1	100	0.863	0.887	0.931	0.877	0.931	0.915	0.933	0.965	0.939	0.977	0.974	0.980	0.995	0.981	0.995
4	1	200	0.877	0.888	0.919	0.884	0.914	0.940	0.936	0.959	0.927	0.959	0.983	0.988	0.996	0.988	0.995
4	1	300	0.889	0.893	0.917	0.896	0.917	0.949	0.950	0.970	0.938	0.963	0.988	0.994	0.997	0.991	0.995
4	1	500	0.887	0.905	0.932	0.891	0.920	0.945	0.954	0.966	0.947	0.967	0.980	0.993	0.995	0.992	0.996
4	1	1000	0.880	0.906	0.921	0.899	0.912	0.931	0.957	0.963	0.949	0.960	0.983	0.985	0.989	0.985	0.990
5	1	20	0.749	0.811	0.915	0.860	0.916	0.843	0.883	0.952	0.915	0.948	0.934	0.954	0.987	0.967	0.986
5	1	50	0.812	0.853	0.946	0.852	0.930	0.884	0.916	0.974	0.917	0.968	0.952	0.976	0.994	0.981	0.995
5	1	100	0.831	0.880	0.945	0.890	0.948	0.909	0.931	0.979	0.942	0.980	0.967	0.983	0.997	0.982	0.998
5	1	200	0.868	0.912	0.958	0.899	0.952	0.920	0.955	0.978	0.947	0.981	0.975	0.990	1.000	0.986	0.995
5	1	300	0.876	0.902	0.951	0.892	0.949	0.924	0.958	0.979	0.953	0.980	0.987	0.989	0.997	0.989	0.998
5	1	500	0.874	0.893	0.937	0.863	0.912	0.942	0.950	0.972	0.922	0.965	0.984	0.986	0.992	0.981	0.995
5	1	1000	0.886	0.893	0.926	0.880	0.914	0.940	0.947	0.967	0.936	0.964	0.988	0.993	0.996	0.983	0.993
7	1	20	0.769	0.790	0.910	0.855	0.916	0.837	0.863	0.948	0.927	0.962	0.914	0.943	0.990	0.987	0.995
7	1	50	0.770	0.859	0.957	0.862	0.933	0.858	0.911	0.982	0.924	0.975	0.957	0.976	0.999	0.986	0.998
7	1	100	0.804	0.872	0.968	0.864	0.942	0.872	0.923	0.986	0.928	0.976	0.958	0.983	1.000	0.985	0.999
7	1	200	0.852	0.903	0.978	0.877	0.956	0.910	0.949	0.991	0.933	0.981	0.977	0.990	0.998	0.983	0.997
7	1	300	0.845	0.901	0.965	0.868	0.962	0.909	0.948	0.985	0.938	0.985	0.980	0.985	0.998	0.986	0.998
7	1	500	0.875	0.882	0.959	0.856	0.932	0.925	0.953	0.984	0.909	0.971	0.985	0.990	0.998	0.976	0.997
7	1	1000	0.863	0.897	0.951	0.836	0.914	0.915	0.949	0.981	0.901	0.961	0.980	0.989	0.998	0.971	0.997

Table C.3: Coverages of Estimated Confidence Intervals for the Aggregate Malmquist Productivity Indices, with $\delta = 0.10$

			0.90					0.95					0.99				
p	q	n	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
1	1	20	0.880	0.880	0.881	0.869	0.870	0.930	0.930	0.929	0.920	0.918	0.982	0.980	0.980	0.974	0.976
1	1	50	0.891	0.891	0.890	0.887	0.888	0.949	0.947	0.948	0.948	0.946	0.992	0.992	0.992	0.991	0.991
1	1	100	0.897	0.900	0.900	0.899	0.897	0.948	0.946	0.946	0.947	0.947	0.989	0.989	0.989	0.989	0.989
1	1	200	0.904	0.903	0.903	0.901	0.901	0.951	0.950	0.950	0.951	0.951	0.988	0.988	0.988	0.988	0.988
1	1	300	0.913	0.916	0.916	0.915	0.915	0.954	0.955	0.955	0.955	0.955	0.992	0.992	0.992	0.992	0.992
1	1	500	0.897	0.895	0.895	0.895	0.895	0.946	0.945	0.945	0.945	0.945	0.988	0.988	0.988	0.988	0.988
1	1	1000	0.906	0.907	0.906	0.907	0.906	0.958	0.959	0.959	0.959	0.959	0.995	0.995	0.995	0.995	0.995
2	1	20	0.824	0.796	0.841	0.790	0.824	0.899	0.865	0.908	0.856	0.889	0.963	0.961	0.975	0.939	0.951
2	1	50	0.869	0.866	0.887	0.830	0.858	0.929	0.928	0.942	0.898	0.922	0.984	0.980	0.987	0.966	0.976
2	1	100	0.875	0.874	0.888	0.845	0.863	0.934	0.926	0.937	0.917	0.923	0.993	0.989	0.991	0.968	0.974
2	1	200	0.907	0.899	0.908	0.873	0.879	0.955	0.951	0.952	0.934	0.936	0.988	0.989	0.989	0.981	0.986
2	1	300	0.898	0.898	0.900	0.897	0.900	0.951	0.952	0.952	0.939	0.944	0.987	0.991	0.992	0.986	0.986
2	1	500	0.898	0.888	0.892	0.885	0.887	0.950	0.945	0.947	0.940	0.944	0.989	0.988	0.988	0.988	0.988
2	1	1000	0.907	0.904	0.907	0.906	0.907	0.953	0.954	0.954	0.951	0.952	0.991	0.991	0.991	0.990	0.991
3	1	20	0.835	0.828	0.897	0.836	0.900	0.909	0.894	0.956	0.906	0.941	0.962	0.964	0.986	0.965	0.988
3	1	50	0.858	0.868	0.921	0.838	0.885	0.930	0.939	0.968	0.904	0.945	0.983	0.983	0.991	0.971	0.992
3	1	100	0.867	0.882	0.910	0.828	0.871	0.924	0.943	0.954	0.893	0.925	0.980	0.981	0.989	0.968	0.982
3	1	200	0.879	0.889	0.907	0.841	0.862	0.937	0.947	0.965	0.907	0.923	0.994	0.992	0.994	0.981	0.989
3	1	300	0.886	0.884	0.894	0.865	0.885	0.941	0.938	0.948	0.924	0.938	0.984	0.987	0.990	0.978	0.987
3	1	500	0.902	0.916	0.929	0.872	0.882	0.957	0.961	0.963	0.934	0.939	0.991	0.991	0.993	0.979	0.981
3	1	1000	0.899	0.903	0.907	0.880	0.889	0.950	0.946	0.948	0.942	0.944	0.993	0.990	0.990	0.983	0.985

Table C.3: Coverages of Estimated Confidence Intervals for the Aggregate Malmquist Productivity Indices, with $\delta = 0.10$ (continued)

			0.90					0.95					0.99				
p	q	n	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)	(i)	(ii)	(iii)	(iv)	(v)
4	1	20	0.782	0.774	0.914	0.833	0.913	0.862	0.860	0.967	0.908	0.954	0.943	0.958	0.995	0.975	0.992
4	1	50	0.844	0.863	0.940	0.881	0.928	0.911	0.923	0.976	0.930	0.965	0.966	0.980	0.998	0.975	0.992
4	1	100	0.872	0.896	0.948	0.866	0.928	0.927	0.948	0.983	0.925	0.975	0.973	0.989	0.996	0.985	0.996
4	1	200	0.874	0.880	0.929	0.856	0.903	0.942	0.942	0.963	0.913	0.954	0.986	0.989	0.997	0.981	0.995
4	1	300	0.887	0.888	0.920	0.842	0.882	0.934	0.938	0.962	0.902	0.940	0.991	0.992	0.996	0.976	0.992
4	1	500	0.893	0.901	0.920	0.866	0.891	0.939	0.945	0.967	0.923	0.936	0.987	0.989	0.994	0.979	0.992
4	1	1000	0.891	0.907	0.921	0.880	0.899	0.944	0.952	0.960	0.936	0.948	0.991	0.983	0.989	0.984	0.987
5	1	20	0.762	0.778	0.922	0.853	0.911	0.840	0.867	0.957	0.909	0.944	0.927	0.950	0.988	0.964	0.988
5	1	50	0.815	0.854	0.954	0.864	0.932	0.869	0.916	0.980	0.921	0.971	0.957	0.976	0.996	0.980	0.996
5	1	100	0.831	0.879	0.959	0.882	0.951	0.904	0.939	0.984	0.938	0.980	0.972	0.989	0.998	0.988	0.999
5	1	200	0.858	0.902	0.962	0.869	0.944	0.920	0.958	0.983	0.940	0.980	0.978	0.991	0.999	0.983	0.993
5	1	300	0.872	0.891	0.947	0.847	0.928	0.936	0.949	0.974	0.926	0.966	0.982	0.990	0.997	0.981	0.994
5	1	500	0.881	0.902	0.938	0.821	0.881	0.932	0.952	0.971	0.897	0.934	0.982	0.989	0.994	0.961	0.987
5	1	1000	0.872	0.907	0.930	0.806	0.856	0.946	0.952	0.972	0.894	0.925	0.989	0.992	0.995	0.972	0.984
7	1	20	0.721	0.761	0.909	0.861	0.906	0.802	0.834	0.943	0.912	0.956	0.902	0.927	0.980	0.980	0.994
7	1	50	0.743	0.847	0.964	0.851	0.939	0.824	0.919	0.987	0.914	0.974	0.932	0.978	0.999	0.985	0.998
7	1	100	0.755	0.869	0.975	0.862	0.948	0.848	0.931	0.992	0.925	0.984	0.934	0.984	0.999	0.986	0.999
7	1	200	0.782	0.893	0.975	0.879	0.949	0.857	0.952	0.993	0.927	0.980	0.942	0.992	0.998	0.983	0.998
7	1	300	0.764	0.909	0.960	0.863	0.961	0.844	0.946	0.985	0.931	0.984	0.946	0.986	0.999	0.986	0.998
7	1	500	0.765	0.885	0.958	0.826	0.910	0.842	0.947	0.985	0.882	0.964	0.943	0.992	1.000	0.968	0.995
7	1	1000	0.754	0.895	0.942	0.782	0.864	0.843	0.942	0.979	0.859	0.938	0.946	0.989	0.997	0.957	0.988