European Journal of Operational Research 277 (2019) 756–769

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

Central limit theorems and inference for sources of productivity change measured by nonparametric Malmquist indices

Léopold Simar[a], Paul W. Wilson[b,*]

[a] Institut de Statistique, Biostatistique, et Sciences Actuarielles, Université Catholique de Louvain-la-Neuve, Louvain-la-Neuve, Belgium
[b] Department of Economics and School of Computing, Division of Computer Science, Clemson University, Clemson, SC, USA

ARTICLE INFO

Article history:
Received 28 August 2018
Accepted 19 February 2019
Available online 25 February 2019

Keywords:
Asymptotic
DEA
Hypothesis test
Inference
Malmquist index

ABSTRACT

Malmquist indices are often used to measure productivity changes in dynamic settings and have been widely applied. The indices are typically estimated using data envelopment analysis (DEA) estimators. Malmquist indices are often decomposed into sub-indices that measure the sources of productivity change (e.g., changes in efficiency, technology or other factors). Recently, Kneip et al. (2018) provide new theoretical results enabling inference about productivity change for individual firms as well as average productivity changed measured in terms of geometric means. This paper extends those results to components of productivity change arising from various decompositions of Malmquist indices. New central limit theorems are developed to allow inference about arithmetic means of logarithms of the sub-indices as well as geometric means of (untransformed) sub-indices. The results are quite general and extend to other sub-indices not explicitly considered in this paper.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

Nonparametric Malmquist indices are widely used to assess changes in productivity across firms in various industries. In addition, Malmquist indices are often decomposed into various measures of sources of productivity change, including changes in efficiency, changes in technology, and other factors. Examples include Aly, Grabowski, Pasurka, and Rangan (1990), Färe, Grosskopf, Lindgren, and Roos (1992, 1994a), Färe, Grosskopf, and Norris (1997), Gilbert and Wilson (1998), Simar and Wilson (1998), Wheelock and Wilson (1999), Alam (2001), Armagan, Ozden, and Bekcioglu (2010), Liu (2010), de Castro Lobo, Ozcan, da Silva, Lins, and Fiszman (2010), Andries (2011), Chang, Hsiao, Huang, and Chang (2011), Chowdhury, Wodchis, and Laporte (2011), Ng (2011), Egilmez and McAvoy (2013), Ahn and Min (2014), Bassem (2014), Wu, Cao, and Liu (2014) and Woo, Chung, Chun, Seo, and Hong (2015).[1] Estimates of both Malmquist indices and their component indices are typically reported for individual firms or units, and often results are summarized by reporting geometric means of estimated Malmquist indices and their corresponding component indices. Geometric means, as opposed to arithmetic means, are used to preserve the multiplicative nature of the indices.

Most applied papers that estimate productivity change and its component sources make no attempt at inference. The few that attempt inference either rely on standard Central Limit Theorem (CLT) results or the bootstrap method proposed by Simar and Wilson (1999). As demonstrated below, however, inferences based on standard CLT results is invalid for cases with more than one input and one output for reasons similar to those discussed by Kneip, Simar, and Wilson (2015) in the context of mean efficiency in cross-sectional settings. Moreover, Simar and Wilson (1999) provide only heuristic arguments to develop their bootstrap method and do not provide any theoretical results. Although the simulation evidence provided by Simar and Wilson (1999) suggests that their smooth bootstrap method works well, the approach cannot be justified theoretically in view of the results obtained below.

Until recently, no theoretical results have been available to permit inference about productivity change estimated by Malmquist indices. Kneip, Simar, and Wilson (2018) establish the convergence rate and the existence of a non-degenerate limiting distribution for data envelopment analysis (DEA) estimators of Malmquist indices for individual producers. These results enable use of the subsampling methods of Simar and Wilson (2011) to make inference about the productivity change from one period to another by an individual producer. In addition, Kneip et al. (2018) provide new central limit theorem (CLT) results for geometric means of Malmquist indices as well as arithmetic means of logarithms of Malmquist

* Corresponding author.
E-mail addresses: leopold.simar@uclouvain.be (L. Simar), pww@clemson.edu (P. W. Wilson).

[1] See also Färe, Grosskopf, and Margaritis (2011) for a recent survey on the use of Malmquist indices.

https://doi.org/10.1016/j.ejor.2019.02.040
0377-2217/© 2019 Elsevier B.V. All rights reserved.

indices over samples of producers.[2] This paper extends the results of Kneip et al. (2018) to component indices obtained by various decompositions of Malmquist indices into sources of productivity change. Theoretical results developed below provide convergence rates and existence of non-degenerate limiting distributions for indices measuring change in efficiency, change in technology, etc. for individual producers, enabling use of the subsampling methods presented by Simar and Wilson (2011) to make inference about individual units. In addition, new CLT results are provided to enable inference about geometric means of the component indices as well as arithmetic means of their logarithms. These new CLT results can be used to make inference about average (geometric or arithmetic) changes in components of productivity change. In addition, the new CLT results can be used for hypothesis testing about differences in changes in efficiency, technology or other features between groups of firms along the lines of Kneip, Simar, and Wilson (2016).

The next section develops a nonparametric, statistical model of production in a dynamic context. Various decompositions of Malmquist indices are considered. The Malmquist index and its component indices are defined in terms of hyperbolic distances (as opposed to distances in the input or output directions) in order to ensure that the component indices are well-defined. In Section 3, hyperbolic DEA estimators and their asymptotic properties are discussed. Near the end of Section 3, new results for these estimators needed for components of Malmquist indices are developed. Results for making inference about components of productivity change are presented in Section 4. In Section 4.1, results for inference about change in technology are developed. These results are then extended to other components of productivity change in Section 4.2. An empirical illustration using data from Färe et al. (1992), is presented in Section 5, and conclusions are discussed in Section 6. Additional technical details, as well as proofs of the theorems presented in Sections 3 and 4 appear in Appendix A.

2. A dynamic, nonparametric production process

In order to establish notation, let $x \in \mathbb{R}^p_+$ and $y \in \mathbb{R}^q_+$ be vectors of fixed input and output quantities. Throughout, vectors are assumed to be column-vectors, as opposed to row-vectors. At time t, the set of feasible combinations of inputs and outputs is given by

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

The *technology*, or *efficient frontier* of Ψ^t, is given by

$$\Psi^{t\partial} := \left\{(x, y) \mid (x, y) \in \Psi^t,\ (\gamma x, \gamma^{-1} y) \notin \Psi^t\ \forall\ \gamma \in (0, 1)\right\}. \tag{2.2}$$

Various economic assumptions regarding Ψ^t can be made; the assumptions of Shephard (1970) and Färe (1988) are typical in microeconomic theory of the firm and are used here.

Assumption 2.1. Ψ^t is closed and strictly convex.

Assumption 2.2. $(x, y) \notin \Psi^t$ if $x = 0,\ y \geq 0, y \neq 0$; i.e., all production requires use of some inputs.

Assumption 2.3. For $\widetilde{x} \geq x,\ \widetilde{y} \leq y$, if $(x, y) \in \Psi^t$ then $(\widetilde{x}, y) \in \Psi^t$ and $(x, \widetilde{y}) \in \Psi^t$; i.e., both inputs and outputs are strongly disposable.

Here and throughout, inequalities involving vectors are defined on an element-by-element basis, as is standard. Assumption 2.2 rules out free lunches, while Assumption 2.3 imposes weak monotonicity on the frontier.

The Farrell (1957) output efficiency measure at time t gives the feasible proportionate expansion of output quantities and is defined by

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

This gives a *radial* measure of efficiency since all output quantities are scaled by the same factor λ. The Farrell (1957) input efficiency measure at time t is given by

$$\theta(x, y \mid \Psi^t) := \inf\left\{\theta \mid (\theta x, y) \in \Psi^t\right\} \tag{2.4}$$

and measures efficiency in terms of the amount by which input levels can be scaled downward by the same factor without reducing output levels. Clearly, $\lambda(x, y|\Psi^t) \geq 1$ and $\theta(x, y|\Psi^t) \leq 1$ for all $(x, y) \in \Psi^t$.

An alternative measure of efficiency is the hyperbolic graph measure of efficiency at time t introduced by Färe, Grosskopf, and Lovell (1985), i.e.,

$$\gamma(x, y \mid \Psi^t) := \inf\left\{\gamma > 0 \mid (\gamma x, \gamma^{-1} y) \in \Psi^t\right\}. \tag{2.5}$$

By construction, $\gamma(x, y|\Psi^t) \leq 1$ for $(x, y) \in \Psi^t$. Just as the measures $\theta(x, y|\Psi^t)$ and $\lambda(x, y|\Psi^t)$ provide measures of the *technical efficiency* of a firm operating at a point $(x, y) \in \Psi^t$, so does $\gamma(x, y|\Psi^t)$, but along a hyperbolic path to the frontier of Ψ^t. The measure in (2.5) gives the amount by which input levels can be feasibly, proportionately scaled downward while simultaneously scaling output levels upward by the same proportion.

Next, define the operator $\mathcal{C}(\cdot)$ so that

$$\mathcal{C}(\Psi^t) := \{(x, y) \mid x = a\widetilde{x},\ y = a\widetilde{y} \text{ for some } (\widetilde{x}, \widetilde{y}) \in \Psi^t \text{ and any } a \in \mathbb{R}^1_+\} \tag{2.6}$$

is the convex cone of the set Ψ^t. Note that this is a pointed cone (i.e., $\mathcal{C}(\Psi^t)$ includes $\{(0, 0)\}$). Analogous to (2.2), the frontier of this set is given by

$$\mathcal{C}^{\partial}(\Psi^t) := \{(x, y) \mid (x, y) \in \mathcal{C}(\Psi^t),\ (\gamma x, \gamma^{-1} y) \notin \mathcal{C}(\Psi^t)\ \forall\ \gamma \in (0, 1)\}. \tag{2.7}$$

If $\mathcal{C}(\Psi^t) = \Psi^t$, then the frontier $\Psi^{t\partial}$ at time t exhibits globally constant returns to scale (CRS), although this is ruled out by strict convexity of Ψ^t in Assumption 2.1. Otherwise, $\Psi^t \subset \mathcal{C}(\Psi^t)$ and $\Psi^{t\partial}$ is said to exhibit variable returns to scale (VRS), with returns to scale either increasing, constant, or decreasing depending on the particular region of the frontier.

Now consider a sample $\mathcal{X}_n = \{(X_i^1, Y_i^1),\ (X_i^2, Y_i^2)\}_{i=1}^n$ of input-output combinations for n firms observed in periods $t = 1$ and 2. To simplify notation, define $Z_i^t := (X_i^t, Y_i^t)$ for $t \in \{1, 2\}$. Then the sample $\mathcal{X}_n$ is represented by $\mathcal{X}_n = \{Z_i^1,\ Z_i^2\}_{i=1}^n$. Firm i's change in productivity between periods 1 and 2 is measured by the hyperbolic Malmquist index

$$\mathcal{M}_i := \left(\frac{\gamma\left(Z_i^2 \mid \mathcal{C}(\Psi^1)\right)}{\gamma\left(Z_i^1 \mid \mathcal{C}(\Psi^1)\right)} \times \frac{\gamma\left(Z_i^2 \mid \mathcal{C}(\Psi^2)\right)}{\gamma\left(Z_i^1 \mid \mathcal{C}(\Psi^2)\right)}\right)^{1/2}. \tag{2.8}$$

This is the geometric mean of two ratios, each providing a measure of productivity change, in the first case using the boundary of $\mathcal{C}(\Psi^1)$ as a benchmark, and in the second case using the boundary of $\mathcal{C}(\Psi^2)$ as a benchmark. For firm i, $\mathcal{M}_i >$ (= or <) 1 if productivity increases (remains unchanged or decreases) between periods 1 and 2. As in Kneip et al. (2018), the Malmquist index here is defined in terms of hyperbolic measures as opposed to input- or output-oriented measures to avoid numerical difficulties. Zofio and Lovell (2001), Johnson and McGinnis (2009) and Russell (2018) discuss the advantages of defining Malmquist indices in terms of hyperbolic distances. In particular, use of hyperbolic measures helps

[2] The results obtained by Kneip et al. (2018) make clear that standard CLT results such as the Lindeberg–Feller CLT cannot be used to make inference about means of logs of Malmquist indices.

ensure that all of the components of productivity change defined below are well-defined.

Various decompositions of Malmquist indices have been proposed in attempts to identify the sources of any changes in productivity. Färe et al. (1992) propose the input-oriented analog of

$$\mathcal{M}_i = \underbrace{\left[\frac{\gamma(Z_i^2 \mid \mathcal{C}(\Psi^2))}{\gamma(Z_i^1 \mid \mathcal{C}(\Psi^1))}\right]}_{:=\mathcal{E}_1(Z_i^1,Z_i^2\mid\Psi^1,\Psi^2)} \times \underbrace{\left[\frac{\gamma(Z_i^2 \mid \mathcal{C}(\Psi^1))}{\gamma(Z_i^2 \mid \mathcal{C}(\Psi^2))} \times \frac{\gamma(Z_i^1 \mid \mathcal{C}(\Psi^1))}{\gamma(Z_i^1 \mid \mathcal{C}(\Psi^2))}\right]^{1/2}}_{:=\mathcal{T}_1(Z_i^1,Z_i^2\mid\Psi^1,\Psi^2)}. \tag{2.9}$$

The authors remark (p. 90) that "the quotient outside the bracket measures the change in technical inefficiency and the ratios inside the bracket measure the shift in the frontier between periods" 1 and 2. However, this is true if and only if the technology is one of globally constant returns to scale. Recognizing this, Färe, Grosskopf, Norris, and Zhang (1994b) decompose the output-oriented analog of $\mathcal{E}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ to obtain the output-oriented analog of

$$\mathcal{M}_i = \underbrace{\left[\frac{\gamma(Z_i^2 \mid \Psi^2)}{\gamma(Z_i^1 \mid \Psi^1)}\right]}_{:=\mathcal{E}_2(Z_i^1,Z_i^2\mid\Psi^1,\Psi^2)} \times \underbrace{\left[\frac{\gamma(Z_i^2 \mid \mathcal{C}(\Psi^2))/\gamma(Z_i^2 \mid \Psi^2)}{\gamma(Z_i^1 \mid \mathcal{C}(\Psi^1))/\gamma(Z_i^1 \mid \Psi^1)}\right]^{1/2}}_{:=\mathcal{S}_1(Z_i^1,Z_i^2\mid\Psi^1,\Psi^2)} \times \mathcal{T}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2). \tag{2.10}$$

Here, $\mathcal{E}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ gives a measure of technical efficiency change under *either* variable or constant returns to scale since efficiency is measured in terms of Ψ^1 and Ψ^2 as opposed to the conical hulls of Ψ^1 and Ψ^2 as in $\mathcal{E}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ in (2.9). The term $\mathcal{S}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ provides a measure of any change in the scale efficiency of firm i. The ratio in the denominator of $\mathcal{S}_1$ measures the distance between the projection of (Z_i^1) onto $\Psi^{1\partial}$ and the projection of (Z_i^1) onto $\mathcal{C}^\partial(\Psi^1)$, providing a measure of the scale efficiency of firm i in period 1.[3] The ratio in the numerator of $\mathcal{S}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ provides the corresponding measure in period 2. It is easy to see that both the numerator and the denominator of $\mathcal{S}_1$ must be less than 1, and that $\mathcal{S}_1 > (=, <)\ 1$ iff scale efficiency for firm i increases (remains unchanged, decreases) from period 1 to period 2.

Both decompositions in (2.9) and (2.10) use the term $\mathcal{T}_1$ to measure change in technology, but this term is based on the conical hulls of Ψ^1 and Ψ^2. Under variable returns to scale, it is possible for the conical hulls to remain unchanged while the technology shifts upward or downward in regions where the technology $\Psi^{t\partial}$ is not coincident with $\mathcal{C}^\partial(\Psi^t)$. This problem is addressed by Ray and Desli (1997) who propose the output-oriented analog of the decomposition

$$\mathcal{M}_i = \mathcal{E}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2) \times \underbrace{\left[\frac{\gamma(Z_i^2 \mid \Psi^1)}{\gamma(Z_i^2 \mid \Psi^2)} \times \frac{\gamma(Z_i^1 \mid \Psi^1)}{\gamma(Z_i^1 \mid \Psi^2)}\right]^{1/2}}_{:=\mathcal{T}_2(Z_i^1,Z_i^2\mid\Psi^1,\Psi^2)} = \underbrace{\left[\frac{\gamma(Z_i^2 \mid \mathcal{C}(\Psi^1))/\gamma(Z_i^2 \mid \Psi^1)}{\gamma(Z_i^1 \mid \mathcal{C}(\Psi^2))/\gamma(Z_i^1 \mid \Psi^2)} \times \frac{\gamma(Z_i^2 \mid \mathcal{C}(\Psi^2))/\gamma(Z_i^2 \mid \Psi^2)}{\gamma(Z_i^1 \mid \mathcal{C}(\Psi^1))/\gamma(Z_i^1 \mid \Psi^1)}\right]^{1/2}}_{:=\mathcal{S}_2(Z_i^1,Z_i^2\mid\Psi^1,\Psi^2)} \tag{2.11}$$

The term $\mathcal{T}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ provides a measure of the change in technology between periods 1 and 2 regardless of whether returns to scale are constant or variable. This term consists of a geometric mean of two ratios. The first ratio gives a measure of any shift in the technology Ψ^∂ relative to firm i's position in period 2. Similarly, the second ratio gives a measure of any shift in the technology relative to firm i's position in period 1. Either of these ratios is greater than (equal to, less than) 1 iff the technology shifts outward (remains unchanged, shifts inward).

Ray and Desli (1997) remark (p. 1036) that $\mathcal{S}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ "is a geometric mean of the ratios of scale efficiencies of the two bundles using in turn the VRS technologies from the two periods as the benchmark. In that sense, it is more in the spirit of a Fisher index." Färe et al. (1997, p. 1042) criticize the measure $\mathcal{S}_2$, and in particular note that the term "may incorrectly identify the scale properties of the underlying technology." while providing an illustrative example in their footnote 7.

Indeed, the term $\mathcal{S}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ defined by (2.11) can be written as

$$\mathcal{S}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2) = \left[\frac{\gamma(Z_i^2 \mid \mathcal{C}(\Psi^1))/\gamma(Z_i^2 \mid \Psi^1)}{\gamma(Z_i^1 \mid \mathcal{C}(\Psi^2))/\gamma(Z_i^1 \mid \Psi^2)} \times \mathcal{S}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)\right]^{1/2}. \tag{2.12}$$

The meaning of $\mathcal{S}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ is clear and intuitive, but the first ratio inside the parentheses in (2.12) is less so. The numerator of this ratio measures scale efficiency in period 1, but from the viewpoint of the firm's location in period 2. Similarly, the denominator measures scale efficiency in period 2, but relative to the firm's location in period 1. Lovell (2003, p. 442) describes $\mathcal{S}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ and notes that "the qualifier 'change' refers to the quantity vectors but not to the technologies."

Gilbert and Wilson (1998), Simar and Wilson (1998) and Wheelock and Wilson (1999) use the output-oriented analog of

$$\mathcal{M}_i = \mathcal{E}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2) \times \mathcal{T}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2) \times \mathcal{S}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2) \times \underbrace{\left[\frac{\gamma(Z_i^1 \mid \mathcal{C}(\Psi^1))/\gamma(Z_i^1 \mid \Psi^1)}{\gamma(Z_i^1 \mid \mathcal{C}(\Psi^2))/\gamma(Z_i^1 \mid \Psi^2)} \times \frac{\gamma(Z_i^2 \mid \mathcal{C}(\Psi^1))/\gamma(Z_i^2 \mid \Psi^1)}{\gamma(Z_i^2 \mid \mathcal{C}(\Psi^2))/\gamma(Z_i^2 \mid \Psi^2)}\right]^{1/2}}_{:=\mathcal{S}_3(Z_i^1,Z_i^2\mid\Psi^1,\Psi^2)} \tag{2.13}$$

after decomposing $\mathcal{S}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ in (2.11) into $\mathcal{S}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ and $\mathcal{S}_3(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$. This measure consists of a geometric mean of two ratios, each resembling the ratio that defines $\mathcal{S}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ in (2.10), but with some important differences. Note that $\mathcal{S}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ measures the change in scale efficiency of *the firm*. This could improve if the firm moves closer to the most efficient scale size in period 2, or it could improve if the firm does not move between periods 1 and 2, but the technology changes so that $\Psi^{2\partial}$ is closer to $\mathcal{C}^\partial(\Psi^2)$ than $\Psi^{1\partial}$ is to $\mathcal{C}^\partial(\Psi^1)$. But now consider the first ratio in the definition of $\mathcal{S}_3(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ in (2.13). Here, the firm's position is fixed at its location in period 1; the ratio can differ from 1 iff the distance between the projection of (Z_i^1) onto $\Psi^{1\partial}$ and $\mathcal{C}^\partial(\Psi^1)$ is different from the projection of (Z_i^1) onto $\Psi^{2\partial}$ and $\mathcal{C}^\partial(\Psi^2)$ along the hyperbolic path through (Z_i^1). The second ratio in $\mathcal{S}_3(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ provides a similar measure relative to the firm's position in period 2, and $\mathcal{S}_3(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ is the geometric mean of these two measures.[4] Gilbert and Wilson (1998), Simar and Wilson (1998) and Wheelock and Wilson (1999) label their corresponding measures "ΔScaleTech," and both Simar and Wilson (1998) and Wheelock

[3] Note that firm i would be scale-efficient in period 1 if $\gamma(Z_i^1 \mid \mathcal{C}(\Psi^1)) = \gamma(Z_i^1 \mid \Psi^1)$. Otherwise, the firm is scale-inefficient. See Wheelock and Wilson (1999) for discussion.

[4] Balk (2001, p. 173) remarks that the decomposition in (2.13) is confusing, but its meaning seems clear.

and Wilson (1999) refer to the term as a measure of "the scale of the technology" as opposed to *the change in the scale efficiency of a firm* as measured by $\mathcal{S}_1$. See Simar and Wilson (1998) and Wheelock and Wilson (1999) for further discussion and illustrations.[5]

All of the quantities and model features defined so far are unobservable, and hence must be estimated. In addition, inference is needed in order to know what might be learned from data. Some additional assumptions are needed to complete the statistical model. The following assumptions are analogous to Assumptions 3.1–3.4 of Kneip et al. (2015). In order to draw upon previous results, we state the assumptions below in terms of the input-oriented measure of efficiency. The assumptions can also be stated in terms of the output, hyperbolic and directional measures of efficiency, and the results of Kneip et al. (2015) extend to those measures after trivial (but tedious) changes in notation in Kneip et al. (2015).

Assumption 2.4. (i) The random variables (X, Y) possess a joint density f^t with support $\mathcal{D}^t \subset \Psi^t$; and (ii) f^t is continuously differentiable on $\mathcal{D}^t$.

Assumption 2.5. (i) $\mathcal{D}^{t*} := \{\theta(x, y \mid \Psi^t)x,\ y) \mid (x, y) \in \mathcal{D}^t\} \subset \mathcal{D}^t$; (ii) $\mathcal{D}^{t*}$ is compact; and (iii) $f^t(\theta(x, y)x, y) > 0$ for all $(x, y) \in \mathcal{D}^t$.

The next two assumptions are needed when DEA estimators are used. Assumption 2.6 imposes some smoothness on the frontier. Kneip, Simar, and Wilson (2008) require only two-times differentiability to establish the existence of a limiting distribution for VRS-DEA estimators, but the stronger assumption that follows is needed to establish results on moments of the DEA estimators.

Assumption 2.6. $\theta(x,\ y|\Psi^t)$ is three times continuously differentiable on $\mathcal{D}^t$.

Recalling that the strong (i.e., free) disposability assumed in Assumption 2.3 implies that the frontier is weakly monotone, the next assumption strengthens this by requiring the frontier to be strictly monotone with no constant segments. This is also needed to establish properties of moments of the DEA estimators.

Assumption 2.7. $\mathcal{D}^t$ is *almost strictly convex*; i.e., for any $(x, y), (\widetilde{x}, \widetilde{y}) \in \mathcal{D}^t$ with $(\frac{x}{\|x\|}, y) \neq (\frac{\widetilde{x}}{\|\widetilde{x}\|}, \widetilde{y})$, the set $\{(x^*, y^*) \mid (x^*, y^*) = (x, y) + \alpha((\widetilde{x}, \widetilde{y}) - (x, y))$ for some $0 < \alpha < 1\}$ is a subset of the *interior* of $\mathcal{D}^t$.

Assumptions 2.1–2.7 comprise a statistical model similar to the one defined in Kneip et al. (2015) and where DEA estimators have desirable properties. However, two additional, important assumptions are needed to obtain asymptotic properties of DEA estimators derived by Kneip et al. (2018) of the Malmquist index defined in (2.8) as well as of DEA estimators of the various components of Malmquist indices presented above. These assumptions appear as Assumptions 3.1 and 3.2 in Kneip et al. (2018). Since these assumptions involve considerable technical detail and require additional notation, the assumptions are presented with some discussion in Section A.1 of Appendix A. Assumption A.1 is needed to ensure well-defined estimators of $\theta(x, y \mid \mathcal{C}(\Psi^t))$ and $\gamma(x, y \mid \mathcal{C}(\Psi^t))$. Part (iii) of the assumption is needed to bound the logarithms of these as well as of $\theta(x,\ y|\Psi^t)$ and $\gamma(x,\ y|\Psi^t)$ away from zero. Assumption A.2 is required to ensure well-defined estimators of cross-period efficiencies.

[5] Other decompositions are possible, and it is not feasible to give an exhaustive treatment here. See Lovell (2003) and Zofio (2007) for summaries and discussion. Note that Lovell (2003) refers to $\mathcal{S}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ as "the activity effect" and decomposes (p. 446) the term into 3 components. Estimation of each of these new components requires nesting one estimator inside another, resulting in considerable complication for statistical inference requiring new theoretical results that are beyond the scope of this paper.

3. Hyperbolic DEA estimators and their asymptotic properties

The VRS-DEA estimator of Ψ^t proposed by Farrell (1957) and Banker, Charnes, and Cooper (1984) is the convex hull of the free-disposal hull of observed input-output pairs in period t. The estimator is given by

$$\widehat{\Psi}_n^t := \left\{(x, y) \in \mathbb{R}^{p+q} \mid y \leq \mathbf{Y}^t\boldsymbol{\omega},\ x \geq \mathbf{X}^t\boldsymbol{\omega},\ \mathbf{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\}, \tag{3.1}$$

where $\mathbf{X}^t = -PLX\,pmatrix - \left(X_1^t, \ldots, X_n^t\right) - PLX\,pmatrix-$ and $\mathbf{Y}^t = -PLX\,pmatrix - \left(Y_1^t, \ldots, Y_n^t\right) - PLX\,pmatrix-$ are $(p \times n)$ and $(q \times n)$ matrices of input and output vectors in period t, respectively; $\mathbf{i}_n$ is an $(n \times 1)$ vector of ones, and $\boldsymbol{\omega}$ is a $(n \times 1)$ vector of weights. Replacing Ψ^t in (2.4) with $\widehat{\Psi}_n^t$ yields the linear program

$$\theta(x, y \mid \widehat{\Psi}_n^t) = \min_{\theta, \boldsymbol{\omega}} \left\{\theta \mid y \leq \mathbf{Y}^t\boldsymbol{\omega},\ \theta x \geq \mathbf{X}^t\boldsymbol{\omega},\ \mathbf{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\}. \tag{3.2}$$

Alternatively, replacing Ψ^t in (2.5) with $\widehat{\Psi}_n^t$ yields the *nonlinear* program

$$\gamma(x, y \mid \widehat{\Psi}_n^t) = \min_{\gamma, \boldsymbol{\omega}} \left\{\gamma \mid \gamma^{-1}y \leq \mathbf{Y}^t\boldsymbol{\omega},\ \gamma x \geq \mathbf{X}^t\boldsymbol{\omega},\ \mathbf{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\}. \tag{3.3}$$

Wilson (2011) provides a simple numerical algorithm for computing $\gamma(x, y \mid \widehat{\Psi}_n^t)$ that avoids the computational difficulty of solving the nonlinear program directly.

Alternatively, the conical DEA (CDEA) estimator $\mathcal{C}(\widehat{\Psi}^t)$ of $\mathcal{C}(\Psi^t)$ is obtained by dropping the constraint $\mathbf{i}_n'\boldsymbol{\omega} = 1$ in (3.1). This leads to the CDEA estimator $\gamma(x, y \mid \mathcal{C}(\widehat{\Psi}^t))$ of $\gamma(x, y \mid \mathcal{C}(\Psi^t))$ obtained by dropping the constraint $\mathbf{i}_n'\boldsymbol{\omega} = 1$ in (3.3).

Kneip et al. (2018) establish asymptotic properties of the CDEA estimator $\gamma(x, y \mid \mathcal{C}(\widehat{\Psi}^t))$ of $\gamma(x, y \mid \mathcal{C}(\Psi^t))$ under appropriate assumptions. In particular, Kneip et al. (2018) establish consistency and existence of a non-degenerate limiting distribution with rate of convergence n^κ under Assumptions 2.1–2.7 where

$$\kappa := \frac{2}{p + q + 1}. \tag{3.4}$$

In addition, Kneip et al. (2018) establish properties of the first two moments of $\gamma(x, y \mid \mathcal{C}(\widehat{\Psi}^t))$ as well as of $\log \gamma(x, y \mid \mathcal{C}(\widehat{\Psi}^t))$.

Kneip et al. (2018) then consider a firm operating at observed, fixed points (x^1, y^1) and (x^2, y^2) in periods 1 and 2. From (2.8) the Malmquist index for this firm is

$$\mathcal{M} = \left[\frac{\gamma\left(x^2, y^2 \mid \mathcal{C}(\Psi^1)\right)}{\gamma\left(x^1, y^1 \mid \mathcal{C}(\Psi^1)\right)} \times \frac{\gamma\left(x^2, y^2 \mid \mathcal{C}(\Psi^2)\right)}{\gamma\left(x^1, y^1 \mid \mathcal{C}(\Psi^2)\right)}\right]^{1/2}. \tag{3.5}$$

Using the data $\mathcal{X}_{n_1}^1 := \{(X_i^1, Y_i^1\}_{i=1,\ldots,n_1}$ and $\mathcal{X}_{n_2}^2 := \{(X_i^2, Y_i^2\}_{i=1,\ldots,n_2}$, $\mathcal{M}$ can be estimated by

$$\widehat{\mathcal{M}} = \left[\frac{\gamma\left(x^2, y^2 \mid \mathcal{C}(\widehat{\Psi}_{n_1}^1)\right)}{\gamma\left(x^1, y^1 \mid \mathcal{C}(\widehat{\Psi}_{n_1}^1)\right)} \times \frac{\gamma\left(x^2, y^2 \mid \mathcal{C}(\widehat{\Psi}_{n_2}^2)\right)}{\gamma\left(x^1, y^1 \mid \mathcal{C}(\widehat{\Psi}_{n_2}^2)\right)}\right]^{1/2}. \tag{3.6}$$

Under Assumptions 2.1–2.7, A.1 and A.2, Theorem 3.3 of Kneip et al. (2018) establishes the existence of a non-degenerate limiting distribution as well as the convergence rate for the estimator in (3.6) of the Malmquist index for a given firm observed in periods 1 and 2. These results permit inference about the unobserved, true Malmquist index $\mathcal{M}$ using the subsampling methods described by Simar and Wilson (2011). In addition, Theorems 4.2 and 4.3 of Kneip et al. (2018) provide CLTs for making inference about $\mu_{\mathcal{M}} :=$

$E(\log \mathcal{M}_i)$ where the expectation is over (X_1, Y_1, X_2, Y_2). In addition, Theorems 4.5 and 4.6 of Kneip et al. (2018) provide CLTs permitting inference about $\exp(\mu_{\mathcal{M}})$ estimated by the geometric mean

$$\widehat{\mathcal{M}}_n := \prod_{i=1}^{n} \left[\frac{\gamma\left(X_i^2, Y_i^2 \mid \mathcal{C}(\widehat{\Psi}_{n_1}^1)\right)}{\gamma\left(X_i^1, Y_i^1 \mid \mathcal{C}(\widehat{\Psi}_{n_1}^1)\right)} \times \frac{\gamma\left(X_i^2, Y_i^2 \mid \mathcal{C}(\widehat{\Psi}_{n_2}^2)\right)}{\gamma\left(X_i^1, Y_i^1 \mid \mathcal{C}(\widehat{\Psi}_{n_2}^2)\right)} \right]^{1/2}. \quad (3.7)$$

Wilson (2011) establishes consistency of the hyperbolic estimator in (3.3) of $\gamma(x, y|\Psi^t)$ under Assumptions weaker than Assumptions 2.1–2.7 listed above, and proves that the rate of convergence is n^{κ}. However, some additional results are needed in order to make inference about the Malmquist index components defined by the various decompositions discussed above in Section 2. Proofs are given in Appendix A.

The first result establishes the existence of non-degenerate limiting distributions for the hyperbolic efficiency estimator and its logarithm.

Theorem 3.1. *Let $\Gamma(\cdot)$ denote either (i) the identity function or (ii) the log function. Under Assumptions 2.1–2.7 and A.1 part (iii),*

$$n^{\kappa}\left(\Gamma\left(\gamma(x, y \mid \widehat{\Psi}_n^t)\right) - \Gamma\left(\gamma(x, y \mid \Psi^t)\right)\right) \xrightarrow{\mathcal{L}} Q_{\gamma}^{\Gamma} \quad (3.8)$$

as $n \to \infty$, where Q_{γ}^{Γ} is a non-degenerate distribution with finite variance.

The next result establishes properties of the first two moments of the hyperbolic efficiency estimator under VRS, analogous to Kneip et al. (2015, Theorem 3.1).

Theorem 3.2. *Let $\Gamma(\cdot)$ denote either (i) the identity function or (ii) the log function. Let $\nu_1 = 3/(p+q+1)$, $\nu_2 = (p+q+4)/(p+q+1)$ and $\nu_3 = (p+q+2)/(p+q+1)$. Under Assumptions 2.1–2.7 and A.1 part (iii), $\exists$ a constant $C_1^{\Gamma} \in (0, \infty)$ such that for all $i, j \in \{1, \ldots, n\}$,*

$$E\left(\Gamma\left(\gamma(X_i^t, Y_i^t \mid \widehat{\Psi}_n^t)\right) - \Gamma\left(\gamma(X_i^t, Y_i^t \mid \Psi^t)\right)\right) = C_1^{\Gamma} n^{-\kappa} + O\left(n^{-\nu_1}(\log n)^{\nu_2}\right), \quad (3.9)$$

$$\mathrm{VAR}\left(\Gamma\left(\gamma(X_i^t, Y_i^t \mid \widehat{\Psi}_n^t)\right) - \Gamma\left(\gamma(X_i^t, Y_i^t \mid \Psi^t)\right)\right) = O\left(n^{-\nu_1}(\log n)^{\nu_1}\right) \quad (3.10)$$

and

$$\left|\mathrm{COV}\left(\Gamma\left(\gamma(X_i^t, Y_i^t \mid \widehat{\Psi}_n^t)\right) - \Gamma\left(\gamma(X_i^t, Y_i^t \mid \Psi^t)\right),\ \Gamma\left(\gamma(X_j^t, Y_j^t \mid \widehat{\Psi}_n^t)\right) - \Gamma\left(\gamma(X_j^t, Y_j^t \mid \Psi^t)\right)\right)\right| = O(n^{-\nu_3}(\log n)^{\nu_3}) = o(n^{-1}). \quad (3.11)$$

The value of the constant C_1^{Γ} depends on the density f, $\Gamma(\cdot)$ and on the structure of the set $\mathcal{D}^t \subset \Psi^t$.

The next result provides properties of moments of the log-hyperbolic estimator in dynamic, two-period settings.

Theorem 3.3. *Let Let ν_1, ν_2 and ν_3 be defined as in Theorem 3.2. Under Assumptions 2.1–2.7, A.1 part (iii) and A.2, for all $t, s \in \{1,2\}$ $\exists$ a constant $C_2^{ts} \in (0, \infty)$ such that for all $i, j \in \{1, \ldots, n\}$,*

$$E\left(\log \gamma(X_i^t, Y_i^t \mid \widehat{\Psi}_n^s) - \log \gamma(X_i^t, Y_i^t \mid \Psi^s)\right) = C_2^{ts} n^{-\kappa} + O\left(n^{-\nu_1}(\log n)^{\nu_2}\right), \quad (3.12)$$

$$\mathrm{VAR}\left(\log \gamma(X_i^t, Y_i^t \mid \widehat{\Psi}_n^s) - \log \gamma(X_i^t, Y_i^t \mid \Psi^s)\right) = O\left(n^{-\nu_1}(\log n)^{\nu_1}\right) \quad (3.13)$$

and for $t^*, s^* \in \{1,2\}$, $j \neq i$,

$$\left|E\left(\left[\log \gamma(X_i^t, Y_i^t \mid \widehat{\Psi}_n^s) - E(\log \gamma(X_i^t, Y_i^t \mid \Psi^s))\right]\left[\log \gamma(X_j^{s^*}, Y_j^{s^*} \mid \widehat{\Psi}_n^{t^*}) - E(\log \gamma(X_j^{s^*}, Y_j^{s^*} \mid \Psi^{t^*}))\right]\right)\right| = O\left(n^{-\nu_3}(\log n)^{\nu_3}\right) = o\left(n^{-1}\right) \quad (3.14)$$

as $n \leq \min\{n_1, n_2\} \to \infty$. The value of the constant C_2^{ts} depends on the density f and on the structure of the sets $\mathcal{D}^s \subset \Psi^s$ and $\mathcal{D}^t \subset \Psi^t$.

4. Inference about Malmquist index components

4.1. Inference about change in technology

This section focuses on the technology change measure $\mathcal{T}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ defined in (2.11) and appearing in (2.13). The measure $\mathcal{T}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ defined in (2.9) is seldom used in the literature, but nonetheless is considered in Section 4.2 as are the other components of productivity change defined in (2.9), (2.10), (2.11) and (2.13).

As discussed above, the sample $\mathcal{X}_n = \{Z_i^1, Z_i^2\}_{i=1}^n$ contains the set of input-output pairs from periods 1 and 2 for firms observed in *both* periods. However, there may be $n_1 > n$ firms observed in period 1, and $n_2 > n$ firms observed in period 2 so that $n \leq \min(n_1, n_2)$. The n_1 observations in $\mathcal{X}_{n_1}^1 = \{Z_i^1\}_{i=1}^{n_1}$ can be used to construct an estimate $\widehat{\Psi}_{n_1}^1$ of Ψ^1, while the n_2 observations in $\mathcal{X}_{n_2}^2 = \{Z_i^2\}_{i=1}^{n_2}$ can be used to construct an estimate $\widehat{\Psi}_{n_2}^2$ of Ψ^2. For a firm observed at $z^1 = (x^1, y^1) \in \Psi^1$ in period 1 and at $z^2 = (x^2, y^2) \in \Psi^2$ in period 2, $\mathcal{T}_2(z^1, z^2 \mid \Psi^1, \Psi^2)$ is estimated by $\mathcal{T}_2(z^1, z^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2)$.

Theorem 4.1. *Under Assumptions 2.1–2.7, A.1 part (iii) and A.2, for each $z^1 \in \mathcal{D}^1$ and $z^2 \in \mathcal{D}^2$, as $n_1, n_2 \to \infty$*

$$n^{\kappa}\left(\mathcal{T}_2(z^1, z^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2) - \mathcal{T}_2(z^1, z^2 \mid \Psi^1, \Psi^2)\right) \xrightarrow{\mathcal{L}} Q_{\mathcal{T}_2, z^1, z^2} \quad (4.1)$$

where $Q_{\mathcal{T}_2, z^1, z^2}$ is a non-degenerate distribution with finite variance.

Theorem 4.1 establishes the existence of a limiting distribution as well as consistency and rate of convergence n^{κ} for the estimator $\mathcal{T}_2(z^1, z^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2)$. These results are sufficient to enable valid inference about $\mathcal{T}_2(z^1, z^2 \mid \Psi^1, \Psi^2)$ for a single firm using the subsampling methods described by Simar and Wilson (2011).

Given the sample $\mathcal{X}_n$, one may obtain n estimates $\mathcal{T}_2(Z_i^1, Z_i^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2)$. Define

$$\begin{aligned}\mu_{\mathcal{T}_2} &:= E\left(\log \mathcal{T}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)\right) \\ &= E(\log \gamma(Z_i^2 \mid \Psi^1) - \log \gamma(Z_i^2 \mid \Psi^2) + \log \gamma(Z_i^1 \mid \Psi^1) \\ &\quad - \log \gamma(Z_i^1 \mid \Psi^2)),\end{aligned} \quad (4.2)$$

where expectations are with respect to Z_i^1 and Z_i^2. Then consider the sample mean

$$\widehat{\mu}_{\mathcal{T}_2, n} := n^{-1} \sum_{i=1}^{n} \log \mathcal{T}_2(Z_i^1, Z_i^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2). \quad (4.3)$$

To simplify notation, let $\sigma_{\mathcal{T}_2}^2 = \mathrm{VAR}(\log \mathcal{T}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)) < \infty$ where expectations are over (Z^1, Z^2). The next result provides a CLT for $\mu_{\mathcal{T}_2}$.

Theorem 4.2. *Under Assumptions 2.1–2.7, A.1 part (iii) and A.2, $\exists$ a constant $D_{\mathcal{T}_2}$ such that*

$$n^{1/2}\left(\widehat{\mu}_{\mathcal{T}_2, n} - \mu_{\mathcal{T}_2} - D_{\mathcal{T}_2} n^{-\kappa} - \xi_{n,\kappa}\right) \xrightarrow{d} N\left(0, \sigma_{\mathcal{T}_2}^2\right) \quad (4.4)$$

where $\xi_{n,\kappa} = O(n^{-\nu_1}(\log n)^{\nu_1}) = o(n^{-\kappa})$ *and* ν_1 *is defined in Theorem 3.2. In addition,*

$$\widehat{\sigma}^2_{\mathcal{T}_2,n} = n^{-1}\sum_{i=1}^{n}\left(\log \mathcal{T}_2(Z_i^1, Z_i^2 \mid \widehat{\Psi}^1_{n_1}, \widehat{\Psi}^2_{n_2}) - \widehat{\mu}_{\mathcal{T}_2,n}\right)^2 \xrightarrow{p} \sigma^2_{\mathcal{T}_2}. \quad (4.5)$$

Although $\widehat{\mu}_{\mathcal{T}_2,n}$ is a consistent estimator of $\mu_{\mathcal{T}_2}$, the estimator has bias $D_{\mathcal{T}_2}n^{-\kappa}$. If $\kappa > 1/2$, then the bias as well as the remainder term $\xi_{n,\kappa}$ are dominated by the $n^{1/2}$ scaling factor and hence can be ignored. Consequently, whenever $\kappa > 1/2$, a $(1-\alpha)\times 100$-percent confidence interval for $\widehat{\mu}_{\mathcal{T}_2,n}$ is estimated by

$$\left[\widehat{\mu}_{\mathcal{T}_2,n} \pm z_{1-\frac{\alpha}{2}}\frac{\widehat{\sigma}_{\mathcal{T}_2,n}}{\sqrt{n}}\right], \quad (4.6)$$

where $z_{1-\frac{\alpha}{2}}$ is the corresponding $(1-\frac{\alpha}{2})$ quantile of the standard normal distribution function. Under the conditions of Theorem 4.2, this interval has asymptotically correct coverage provided $\kappa > 1/2$ (i.e., $p+q \le 2$).

By contrast, if $\kappa = 1/2$, the bias in (4.4) is constant. If $\kappa < 1/2$, the bias tends to infinity as $n \to \infty$. In cases where $\kappa \le 1/2$, replacing the scaling factor $n^{1/2}$ with n^{ζ} where $\zeta \in (0, \kappa)$ would drive the bias to 0 as $n \to \infty$, but would also drive the variance to 0, resulting in a degenerate limiting distribution and preventing inference from begin made. The usefulness of Theorem 4.2 for practical applications is quite limited since $\kappa > 1/2$ if and only if $(p+q) \le 2$. Fortunately, an approach similar to the one of Kneip et al. (2015) can be used to solve this problem.

Let $n_\kappa = \min(\lfloor n^{2\kappa}\rfloor, n)$, where $\lfloor a\rfloor$ denotes the largest integer less than or equal to a. Then for $\kappa < 1/2$, $n_\kappa < n$. Assume that the observations in $\mathcal{X}_n$ are randomly sorted (the algorithm described by Daraio, Simar, and Wilson (2018), Appendix D can be used to randomly sort the observations while allowing results to be replicated by other researchers using the same data and the same sorting algorithm). Define

$$\widehat{\mu}_{\mathcal{T}_2,n_\kappa} := n_\kappa^{-1}\sum_{i=1}^{n_\kappa}\log\left(\mathcal{T}_2(Z_i^1, Z_i^2 \mid \widehat{\Psi}^1_{n_1}, \widehat{\Psi}^2_{n_2})\right). \quad (4.7)$$

Note that the estimates $\mathcal{T}_2(Z_i^1, Z_i^2 \mid \widehat{\Psi}^1_{n_1}, \widehat{\Psi}^2_{n_2})$ are computed using all of the available observations, but that the summation is over only the first n_κ observations in $\mathcal{X}_n$. The next result establishes the properties of this estimator.

Theorem 4.3. *Under Assumptions 2.1–2.7, A.1 part (iii) and A.2, for cases where* $\kappa \le 1/2$,

$$n^{\kappa}\left(\widehat{\mu}_{\mathcal{T}_2,n_\kappa} - \mu_{\mathcal{T}_2} - D_{\mathcal{T}_2}n^{-\kappa} - \xi_{n,\kappa}\right) \xrightarrow{d} N\left(0, \sigma^2_{\mathcal{T}_2}\right) \quad (4.8)$$

as $n \to \infty$, *where* $\xi_{n,\kappa} = O(n^{-\nu_1}(\log n)^{\nu_1}) = o(n^{-\kappa})$ *and* ν_1 *is defined in Theorem 3.2.*

The bias term $D_{\mathcal{T}_2}n^{-\kappa}$ remains in (4.8), but the both the bias and the variance remain constant as $n \to \infty$. Consequently, the bias term can be replaced with a generalized jackknife estimate $\widehat{B}_{\mathcal{T}_2,n,\kappa}$ similar to the bias estimate developed in Kneip et al. (2015). The bias estimate presented in Kneip et al. (2015) assumes $n_1 = n_2 = n$, while the presentation below explicitly allows $n_1 \ge n$ or $n_2 \ge n$.

Recall that n firms are observed in both periods 1 and 2; these observations comprise the sample $\mathcal{X}_n = \{(Z_i^1, Z_i^2)\}$. In addition, as discussed above, there are $n_1^* = n_1 - n \ge 0$ firms observed in period 1 but not in period 2; let these observations comprise the set $\mathcal{W}^1_{n_1^*} = \{W_i^1\}_{i=1}^{n_1^*} \subset \mathcal{X}^1_{n_1}$. Similarly, there are $n_2^* = n_2 - n \ge 0$ firms observed in period 2 but not in period 1; let $\mathcal{W}^2_{n_2^*} = \{W_i^2\}_{i=1}^{n_2^*} \subset \mathcal{X}^2_{n_2}$ denote the set of such observations. Of course, either $\mathcal{W}^1_{n_1^*}$ or $\mathcal{W}^2_{n_2^*}$ will be the empty set if $n_1 = n$ or $n_2 = n$. Now split $\mathcal{X}_n$ randomly into two sub-samples $\mathcal{X}^{(1)}_{m_1}$ and $\mathcal{X}^{(2)}_{m_2}$ of sizes $m_1 = \lfloor n/2\rfloor$ and $m_2 = n - \lfloor n/2\rfloor$ (respectively). Note that if n is even, $m_1 = m_2$, but if n is odd then $m_1 = m_2 - 1$. Asymptotically, this makes no difference since $m_1/m_2 \to 1$ as $n \to \infty$. In addition, split $\mathcal{W}^1_{n_1^*}$ randomly into two sub-samples $\mathcal{W}^{1(1)}_{m_{11}}$ and $\mathcal{W}^{1(2)}_{m_{12}}$ of sizes $m_{11} = \lfloor n_1^*/2\rfloor$ and $m_{12} = n_1^* - \lfloor n_1^*/2\rfloor$ (respectively). If n_1^* is even, $m_{11} = m_{12}$, but if n_1^* is odd then $m_{11} = m_{12} - 1$, but this also makes no difference asymptotically. Similarly, split $\mathcal{W}^2_{n_2^*}$ randomly into two sub-samples $\mathcal{W}^{2(2)}_{m_{21}}$ and $\mathcal{W}^{2(2)}_{m_{22}}$ of sizes $m_{21} = \lfloor n_2^*/2\rfloor$ and $m_{22} = n_1^* - \lfloor n_2^*/2\rfloor$ (respectively). If n_2^* is even, $m_{21} = m_{22}$, but if n_2^* is odd then $m_{21} = m_{22} - 1$, but again this also makes no difference asymptotically.

Now let $\mathcal{X}^{t(j)}_{m_j}$ denote the set of observations on Z_i^t for period $t \in \{1,2\}$ and subsample $j \in \{1,2\}$. Let $m^*_{tj} = m_j + m_{tj}$. Define $\mathcal{V}^{t(j)}_{m^*_{tj}} := \mathcal{X}^{(t)}_{m_j} \cup \mathcal{W}^{t(j)}_{m_{tj}}$. Let $\widehat{\Psi}^{t(j)}_{m^*_{tj}}$ denote the estimator of Ψ^t, analogous to (3.1), but obtained using the observations in $\mathcal{V}^{t(j)}_{m^*_{tj}}$ instead of $\mathcal{X}^t_{n_1}$. Let $\gamma(x, y \mid \widehat{\Psi}^{t(j)}_{m^*_{tj}})$ denote the corresponding estimator of $\gamma(x, y|\Psi)$ obtained by substituting $\widehat{\Psi}^{t(j)}_{m^*_{tj}}$ for Ψ in (2.5).

Now define

$$\widehat{\mu}^{(j)}_{\mathcal{T}_2,m_j} := m_j^{-1}\sum_{i|(Z_i^1,Z_i^2)\in\mathcal{X}^{(j)}_{m_j}}\log\widehat{\mathcal{T}}_2(W_i^1, W_i^2 \mid \widehat{\Psi}^{1(j)}_{m^*_{1j}}, \widehat{\Psi}^{2(j)}_{m^*_{2j}}) \quad (4.9)$$

for $j \in \{1,2\}$ and set

$$\widehat{\mu}^*_{\mathcal{T}_2,n/2} = \frac{1}{2}\left(\widehat{\mu}^{(1)}_{\mathcal{T}_2,m_1} + \widehat{\mu}^{(2)}_{\mathcal{T}_2,m_2}\right). \quad (4.10)$$

Using reasoning similar to that in Kneip et al. (2015, Section 4), it is easy to show that

$$\widetilde{B}_{\mathcal{T}_2,n,\kappa} = (2^\kappa - 1)^{-1}\left(\widehat{\mu}^*_{\mathcal{T}_2,n/2} - \widehat{\mu}_{\mathcal{M},n}\right) = D_{\mathcal{T}_2}n^{-\kappa} + \xi^*_{n,\kappa} + o_p\left(n^{-1/2}\right), \quad (4.11)$$

provides an estimator of the bias $D_{\mathcal{T}_2}n^{-\kappa}$. The remainder term $\xi^*_{n,\kappa}$ in (4.11) is of the same order as $\xi_{n,\kappa}$ appearing in (4.4).

Note that there are $\binom{n}{n/2}$ possible splits of the original n observations. To reduce the variance of the bias estimate in (4.11), the sample can be split $K << \binom{n}{n/2}$ times while randomly shuffling the observations before each split, and computing $\widetilde{B}_{\mathcal{T}_2,n,\kappa,k}$ using (4.11) for $k = 1, \ldots, K$. Then

$$\widehat{B}_{\mathcal{T}_2,n,\kappa} = K^{-1}\sum_{k=1}^{K}\widetilde{B}_{\mathcal{T}_2,n,\kappa,k} \quad (4.12)$$

gives a generalized jackknife estimate of the bias $D_{\mathcal{T}_2}n^{-\kappa}$ (Gray & Schucany, 1972, Definition 2.1). Averaging in (4.12) reduces the variance by a factor of K^{-1} relative to the bias in (4.11).

Substituting the bias estimate in (4.12) for the bias terms $D_{\mathcal{T}_2}$ in Theorems 4.2 and 4.3 leads to the following CLT result.

Theorem 4.4. *Under the conditions of Theorem 4.2, as* $n \to \infty$

$$n^{1/2}\left(\widehat{\mu}_{\mathcal{T}_2,n} - \widehat{B}_{\mathcal{T}_2,n,\kappa} - \mu_{\mathcal{T}_2} - \xi_{n,\kappa}\right) \xrightarrow{d} N\left(0, \sigma^2_{\mathcal{T}_2}\right) \quad (4.13)$$

whenever $\kappa \ge 2/5$. *In addition, for cases where* $\kappa < 1/2$,

$$n^{\kappa}\left(\widehat{\mu}_{\mathcal{T}_2,n_\kappa} - \widehat{B}_{\mathcal{T}_2,n,\kappa} - \mu_{\mathcal{T}_2} - \xi_{n,\kappa}\right) \xrightarrow{d} N\left(0, \sigma^2_{\mathcal{T}_2}\right) \quad (4.14)$$

as $n \to \infty$.

Note that in all cases (i.e., for all values of κ), $\xi_{n,\kappa} = o(n^{-\kappa})$ and hence $n^{\kappa}\xi_{n,\kappa} = o(1)$. Therefore the remainder term can be neglected.

In cases where $\kappa \ge 2/5$ and hence $(p+q) \le 4$, Theorem 4.4 together with (4.5) from Theorem 4.2 ensures that the interval

$$\left[\widehat{\mu}_{\mathcal{T}_2,n} - \widehat{B}_{\mathcal{T}_2,n,\kappa} \pm z_{1-\frac{\alpha}{2}}\frac{\widehat{\sigma}_{\mathcal{T}_2,n}}{\sqrt{n}}\right], \quad (4.15)$$

where as in (4.6) $z_{1-\frac{\alpha}{2}}$ represents the $\left(1-\frac{\alpha}{2}\right)$ quantile of the standard normal distribution function provides an asymptotically correct $(1-\alpha)$ confidence interval for $\mu_{\mathcal{T}_2}$. For cases where $\kappa < 1/2$ and hence $(p+q) \geq 4$, Theorem 4.4 permits construction of the asymptotically correct $(1-\alpha)$ confidence interval

$$\left[\widehat{\mu}_{\mathcal{T}_2,n_\kappa} - \widehat{B}_{\mathcal{T}_2,n,\kappa} \pm z_{1-\frac{\alpha}{2}} \frac{\widehat{\sigma}_{\mathcal{T}_2,n}}{n^\kappa}\right] \tag{4.16}$$

for $\mu_{\mathcal{T}_2}$.

The interval in (4.16) is centered on $\widehat{\mu}_{\mathcal{T}_2,n_\kappa} - \widehat{B}_{\mathcal{T}_2,n,\kappa}$, and $\widehat{\mu}_{\mathcal{T}_2,n_\kappa}$ is computed from a random subset of n_κ estimates $\mathcal{T}_2(Z_i^1, Z_i^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2)$. As discussed by Kneip et al. (2015), while this may seem arbitrary, *any* confidence interval for $\mu_{\mathcal{T}_2}$ is arbitrary since any asymmetric confidence interval for $\mu_{\mathcal{T}_2}$ can be constructed simply by using different quantiles of the $N(0, 1)$ distribution to establish the bounds. The goal is always to achieve a high level of coverage without making the confidence interval too wide to be informative.

Alternatively, in cases where $\kappa < 1/2$, the randomness of the interval in (4.16) due to centering on a mean over a subsample of size $n_\kappa < n$ can be eliminated by replacing $\widehat{\mu}_{\mathcal{T}_2,n_\kappa}$ with $\widehat{\mu}_{\mathcal{T}_2,n}$ to obtain

$$\left[\widehat{\mu}_{\mathcal{T}_2,n} - \widehat{B}_{\mathcal{T}_2,n,\kappa} \pm z_{1-\frac{\alpha}{2}} \frac{\widehat{\sigma}_{\mathcal{T}_2,n}}{n^\kappa}\right]. \tag{4.17}$$

Both intervals (4.16) and (4.17) have the same length and hence are equally informative. However, the interval in (4.17) should have higher coverage in finite samples because the estimator $\widehat{\mu}_{\mathcal{T}_2,n}$ uses more information than the estimator $\widehat{\mu}_{\mathcal{T}_2,n_\kappa}$. Hence for $\kappa < 1/2$, $n_\kappa < n$ and hence the interval in (4.17) contains the true value $\mu_{\mathcal{T}_2}$ with probability greater than $(1-\alpha)$. Due to the results given above, it is clear that the coverage of the interval in (4.17) converges to 1 as $n \to \infty$.

Note that when $(p+q) = 4$, either result (4.13) or (4.14) can be used to construct intervals with asymptotically correct coverage. For reasons given by Kneip et al. (2015, Section 4.1), one should expect the interval in (4.16) to provide a better approximation in finite samples than (4.15) when $(p+q) = 4$.

As with estimates of the Malmquist index defined in (2.8), researchers typically report geometric means

$$\widehat{\mathcal{F}}_{2,n} = \exp\left(\widehat{\mu}_{\mathcal{T}_2,n}\right) = \left(\prod_{i=1}^{n} \widehat{\mathcal{T}}_2(Z_i^1, Z_i^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2)\right)^{1/n}. \tag{4.18}$$

Clearly, $\widehat{\mathcal{F}}_{2,n}$ can be seen as an estimator of $\mathcal{F}_2 = \exp(\mu_{\mathcal{T}_2})$. The properties of this estimator are given the next result.

Theorem 4.5. *Under Assumptions 2.1–2.7, A.1 part (iii) and A.2, as* $n \leq \min\{n_1, n_2\} \to \infty$

$$n^{1/2}\left(\widehat{\mathcal{F}}_{2,n} - \exp(\mu_{\mathcal{T}_2}) + \exp(\mu_{\mathcal{T}_2}) D_{\mathcal{T}_2} n^{-\kappa} + \xi_{\mathcal{F}_2,n,\kappa}\right) \xrightarrow{d} N\left(0,\ \exp(2\mu_{\mathcal{T}_2})\sigma_{\mathcal{T}_2}^2\right) \tag{4.19}$$

where $\xi_{\mathcal{F}_2,n,\kappa} = O(n^{-\nu_1}(\log n)^{\nu_1}) = o(n^{-\kappa})$ *and* ν_1 *is defined in Theorem 3.2.*

Provided $\kappa > 1/2$, both the bias and the remainder terms in (4.19) are asymptotically negligible, and

$$\left[\widehat{\mathcal{F}}_{2,n} \pm z_{1-\frac{\alpha}{2}} \frac{\exp\left(\widehat{\mu}_{\mathcal{T}_2,n}\right)\widehat{\sigma}_{\mathcal{T}_2,n}}{n^{1/2}}\right] \tag{4.20}$$

provides a $(1-\alpha) \times 100$-percent confidence interval for $\exp(\mu_{\mathcal{T}_2})$ with asymptotically correct coverage. But if $\kappa \leq 1/2$, the bias must be dealt with.

Suppose $\kappa \leq 1/2$. Assume the observations in $\mathcal{X}_n$ are randomly ordered and define

$$\widehat{\mathcal{F}}_{2,n_\kappa} = \exp\left(\widehat{\mu}_{\mathcal{T}_2,n_\kappa}\right) = \left(\prod_{i=1}^{n_\kappa} \widehat{\mathcal{T}}_2(Z_i^1, Z_i^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2)\right)^{1/n_\kappa}. \tag{4.21}$$

Note that the estimates under the product sign are computed using all of the available data, but the product is over only the first n_κ observations in $\mathcal{X}_n$. The properties of the estimator $\widehat{\mathcal{F}}_{2,n_\kappa}$ are established in the next theorem.

Theorem 4.6. *Under Assumptions 2.1–2.7, A.1 part (iii) and A.2, for cases where* $\kappa \leq 1/2$,

$$n^\kappa\left(\widehat{\mathcal{F}}_{2,n_\kappa} - \exp(\mu_{\mathcal{T}_2}) + \exp(\mu_{\mathcal{T}_2}) D_{\mathcal{T}_2} n^{-\kappa} + \xi_{\mathcal{F}_2,n,\kappa}\right) \xrightarrow{d} N\left(0,\ \exp(2\mu_{\mathcal{T}_2})\sigma_{\mathcal{T}_2}^2\right) \tag{4.22}$$

as $n \to \infty$.

As in Theorem 4.3, the bias is stabilized in Theorem 4.6, but it does not disappear as $n \to \infty$ and therefore must be estimated. A generalized jackknife estimate $\widehat{B}_{\mathcal{F}_2,n,\kappa}$ analogous to the estimate $\widehat{B}_{\mathcal{T}_2,n,\kappa}$ discussed above can be obtained by following the steps to compute $\widehat{B}_{\mathcal{T}_2,n,\kappa}$ but replacing the sample arithmetic means with their corresponding sample geometric means. This leads to the following result.

Theorem 4.7. *Under the conditions of Theorem 4.2,*

$$n^{1/2}\left(\widehat{\mathcal{F}}_{2,n} - \widehat{B}_{\mathcal{F}_2,n,\kappa} - \exp(\mu_{\mathcal{T}_2}) + \xi_{\mathcal{F}_2,n,\kappa}\right) \xrightarrow{d} N\left(0, \exp(2\mu_{\mathcal{T}_2})\sigma_{\mathcal{T}_2}^2\right) \tag{4.23}$$

as $n \to \infty$ *whenever* $\kappa \geq 2/5$. *In addition, for cases where* $\kappa < 1/2$,

$$n^\kappa\left(\widehat{\mathcal{F}}_{2,n_\kappa} - \widehat{B}_{\mathcal{F}_2,n,\kappa} - \exp(\mu_{\mathcal{T}_2}) - \xi_{\mathcal{F}_2,n,\kappa}\right) \xrightarrow{d} N\left(0, \exp(2\mu_{\mathcal{T}_2})\sigma_{\mathcal{T}_2}^2\right) \tag{4.24}$$

as $n \to \infty$.

For cases where $\kappa \geq 2/5$, Theorem 4.7 permits construction of an asymptotically correct $(1-\alpha)$ confidence interval for $\exp(\mu_{\mathcal{T}_2})$ given by

$$\left[\widehat{\mathcal{F}}_{2,n} - \widehat{B}_{\mathcal{F}_2,n,\kappa} \pm z_{1-\frac{\alpha}{2}} \frac{\exp\left(\widehat{\mu}_{\mathcal{T}_2,n}\right)\widehat{\sigma}_{\mathcal{T}_2,n}}{n^{1/2}}\right]. \tag{4.25}$$

Alternatively, whenever $\kappa < 1/2$, Theorem 4.7 can be used to construct the asymptotically correct $(1-\alpha)$ confidence interval

$$\left[\widehat{\mathcal{F}}_{2,n_\kappa} - \widehat{B}_{\mathcal{F}_2,n,\kappa} \pm z_{1-\frac{\alpha}{2}} \frac{\exp\left(\widehat{\mu}_{\mathcal{T}_2,n}\right)\widehat{\sigma}_{\mathcal{T}_2,n}}{n^\kappa}\right]. \tag{4.26}$$

Analogous to the discussion above, one could also replace $\widehat{\mathcal{F}}_{2,n_\kappa}$ with $\widehat{\mathcal{F}}_n$ in (4.26), with the coverage of the resulting interval converging to 1 as $n \to \infty$.

Also as discussed above, one can use either of the intervals in (4.25) and (4.26) when $(p+q) = 4$. The interval in (4.25) uses the scaling factor $\sqrt{n}$ and hence neglects the term $\sqrt{n}\eta_{n,\kappa} = O(n^{-1/10})$ in result (4.23) of Theorem 4.7, while the interval in (4.26) uses the scaling factor n^κ and hence neglects the term $n^\kappa \eta_{n,\kappa} = O(n^{-1/5})$ in result (4.24) of Theorem 4.7. Therefore one should expect (4.26) to provide a better approximation in finite samples than (4.25) when $(p+q) = 4$. For testing purposes, however, one cannot escape the tradeoff between size and power.

The null hypothesis of no technology change corresponds to $\exp(\mu_{\mathcal{T}_2}) = 1$, while the alternative hypothesis of change in technology between periods 1 and 2 corresponds to $\exp(\mu_{\mathcal{T}_2}) \neq 1$. Hence the null is rejected whenever the relevant estimated confidence interval in (4.25) or (4.26) does not include unity. The results of such tests are expected to be similar to the results of similar tests based on log values, but small differences may arise due to the different asymptotic approximations involved. Asymptotically, any differences are negligible.

4.2. Inference about other components of productivity change

From an applications perspective, the most important results in Section 4.1 are Theorems 4.1, 4.4 and 4.7. The results

in Theorems 4.2 and 4.3 are intermediate results needed to establish Theorem 4.1 and make clear the role of estimation bias. Theorems 4.5 and 4.6 similarly lead to Theorem 4.7. Just as $\mathcal{T}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ is estimated by $\mathcal{T}_2(Z_i^1, Z_i^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2)$ in Section 4.1, each of the components $\mathcal{E}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$, $\mathcal{E}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$, $\mathcal{T}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$, $\mathcal{S}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$, $\mathcal{S}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ and $\mathcal{S}_3(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ defined in (2.9), (2.10), (2.11) and (2.13) can be estimated by replacing Ψ^1 and Ψ^2 in the definitions of the measures by the estimators $\widehat{\Psi}_{n_1}^1$ and $\widehat{\Psi}_{n_2}^2$.

A careful reading of the proofs of Theorems 4.4–4.7 in Appendix A reveals that arguments similar to those used to obtain the results for change in technology in Section 4.1 can be used to establish analogous results for an estimator of the change-in-efficiency measure $\mathcal{E}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$, which like the estimator of $\mathcal{T}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ involves a ratio of measures $\gamma(Z_i^s \mid \Psi^t)$. The other components of productivity change listed above involve ratios of both $\gamma(Z_i^s \mid \Psi^t)$ and $\gamma(Z_i^s \mid \mathcal{C}(\Psi^t))$, $s, t \in \{1,2\}$. Arguments similar to those used in the proofs of the results in Section 4.1, combined with results from Kneip et al. (2018) on the CDEA estimator of distances to boundaries of conical hulls $\mathcal{C}(\Psi^t)$, can be used to derive results for estimators of $\mathcal{E}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$, $\mathcal{T}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$, $\mathcal{S}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$, $\mathcal{S}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ and $\mathcal{S}_3(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ analogous to those obtained in Section 4.1 for the estimator of $\mathcal{T}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$. Consequently, to avoid repetition, the results in this section are stated without formal proofs.

To simplify notation, let Ξ be a place-holder denoting either $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{T}_1$, $\mathcal{S}_1$, $\mathcal{S}_2$ or $\mathcal{S}_3$ or some other index defined in terms of ratios of the measures $\gamma(Z_i^s, Z_i^t \mid \Psi^s, \Psi^t)$ or $\gamma(Z_i^s, Z_i^t \mid \mathcal{C}(\Psi^s), \mathcal{C}(\Psi^t))$, $s, t \in \{1,2\}$. The results that follow hold when Ξ is replaced with any of the components listed above. The next result is immediate.

Theorem 4.8. *Under Assumptions 2.1–2.7, A.1 and A.2, for each $z^1 \in \mathcal{D}^1$ and $z^2 \in \mathcal{D}^2$, as $n_1, n_2 \to \infty$*

$$n^{\kappa}\left(\Xi(z^1, z^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2) - \Xi(z^1, z^2 \mid \Psi^1, \Psi^2)\right) \xrightarrow{\mathcal{L}} Q_{\Xi, z^1, z^2} \quad (4.27)$$

where Q_{Ξ, z^1, z^2} is a non-degenerate distribution with finite variance.

Remark 4.1. Note that in the theorems of Section 4.1, parts (i)–(ii) of Assumption A.1 are not needed since the measure $\mathcal{T}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ does not involve the conical hull of either Ψ^1 or Ψ_2. Similarly, $\mathcal{E}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ also does not involve the conical hull of either Ψ^1 or Ψ^2. Hence when $\mathcal{E}_2$ replaces Ξ in (4.27), Theorem 4.8 does not require parts (i)–(ii) of Assumption A.1. But for the other measures listed above, all three parts of Assumption A.1 are needed. The same remark applies to the remaining theorems that follow in this section.

Theorem 4.8 establishes the existence of limiting distributions as well as consistency and rate of convergence n^{κ} for the estimators of the components $\mathcal{E}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$, $\mathcal{E}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$, $\mathcal{T}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$, $\mathcal{S}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$, $\mathcal{S}_2(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ and $\mathcal{S}_3(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ of productivity change. These results are sufficient to enable valid inference about each component for a single firm using the subsampling methods described by Simar and Wilson (2011).

Analogous to (4.2), define

$$\mu_{\Xi} := E\left(\log \Xi(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)\right) \quad (4.28)$$

and consider the sample mean

$$\widehat{\mu}_{\Xi,n} := n^{-1} \sum_{i=1}^{n} \log \Xi(Z_i^1, Z_i^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2). \quad (4.29)$$

Similar to (4.7), define

$$\widehat{\mu}_{\Xi,n_{\kappa}} := n_{\kappa}^{-1} \sum_{i=1}^{n_{\kappa}} \log \Xi(Z_i^1, Z_i^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2), \quad (4.30)$$

noting that the estimates $\log \Xi(Z_i^1, Z_i^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2)$ are computed using all of the available observations, but that the summation is over only the first n_{κ} observations in $\mathcal{X}_n$ where again $n_{\kappa} = \min(\lfloor n^{2\kappa} \rfloor, n)$. Finally, let $\widehat{B}_{\Xi,n,\kappa}$ denote the generalized jackknife estimate of bias $D_{\Xi} n^{-\kappa}$ analogous to $\widehat{B}_{\mathcal{T}_2,n,\kappa}$ in (4.12) obtained by replacing $\mathcal{T}_2$ with Ξ in (4.9)–(4.12). The next result enables inference about μ_{Ξ}.

Theorem 4.9. *Under the conditions of Theorem 4.2, as $n \to \infty$*

$$n^{1/2}\left(\widehat{\mu}_{\Xi,n} - \widehat{B}_{\Xi,n,\kappa} - \mu_{\Xi} - \xi_{\Xi,n,\kappa}\right) \xrightarrow{d} N\left(0, \sigma_{\Xi}^2\right) \quad (4.31)$$

whenever $\kappa \geq 2/5$. Alternatively, for cases where $\kappa < 1/2$,

$$n^{\kappa}\left(\widehat{\mu}_{\Xi,n_{\kappa}} - \widehat{B}_{\Xi,n,\kappa} - \mu_{\Xi} - \xi_{\Xi,n,\kappa}\right) \xrightarrow{d} N\left(0, \sigma_{\Xi}^2\right) \quad (4.32)$$

as $n \to \infty$. In addition,

$$\widehat{\sigma}_{\Xi,n}^2 = n^{-1} \sum_{i=1}^{n} \left(\log \Xi(Z_i^1, Z_i^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2) - \widehat{\mu}_{\Xi,n}\right)^2 \xrightarrow{p} \sigma_{\Xi}^2. \quad (4.33)$$

In all cases (i.e., for all values of κ), $\xi_{\Xi,n,\kappa} = o(n^{-\kappa})$ and hence $n^{\kappa} \xi_{\Xi,n,\kappa} = o(1)$. Therefore the remainder term can be neglected. Theorem 4.9 ensures that the interval

$$\left[\widehat{\mu}_{\Xi,n} - \widehat{B}_{\Xi,n,\kappa} \pm z_{1-\frac{\alpha}{2}} \frac{\widehat{\sigma}_{\Xi,n}}{\sqrt{n}}\right] \quad (4.34)$$

provides a confidence interval for μ_{Ξ} with asymptotically correct coverage of $(1-\alpha)$ in cases where $\kappa \geq 2/5$. Alternatively, when $\kappa < 1/2$, Theorem 4.9 ensures that the interval

$$\left[\widehat{\mu}_{\Xi,n_{\kappa}} - \widehat{B}_{\Xi,n,\kappa} \pm z_{1-\frac{\alpha}{2}} \frac{\widehat{\sigma}_{\Xi,n}}{n^{\kappa}}\right] \quad (4.35)$$

has asymptotic coverage of $(1-\alpha)$.

In order to consider geometric means of the various components of productivity change while avoiding repetitive notation, let $\mathscr{E}_1$, $\mathscr{E}_2$, $\mathscr{T}_1$, $\mathscr{T}_2$, $\mathscr{S}_1$, $\mathscr{S}_2$ or $\mathscr{S}_3$ denote geometric means of estimators of $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{T}_1$, $\mathcal{T}_2$, $\mathcal{S}_1$, $\mathcal{S}_2$ or $\mathcal{S}_3$, respectively. In other words, write

$$\widehat{\Upsilon}_n = \exp\left(\widehat{\mu}_{\Xi,n}\right) = \left(\prod_{i=1}^{n} \widehat{\Xi}(Z_i^1, Z_i^2 \mid \widehat{\Psi}_{n_1}^1, \widehat{\Psi}_{n_2}^2)\right)^{1/n} \quad (4.36)$$

and replace Ξ with one of $\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{T}_1, \mathcal{T}_2, \mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3\}$ while replacing Υ with the corresponding element of the set $\{\mathscr{E}_1, \mathscr{E}_2, \mathscr{T}_1, \mathscr{T}_2, \mathscr{S}_1, \mathscr{S}_2, \mathscr{S}_3\}$ (for example, replacing Ξ and Υ in (4.36) with $\mathcal{T}_2$ and $\mathscr{T}_2$, respectively, yields the expression in (4.18)). Then let $\widehat{B}_{\Upsilon,n,\kappa}$ denote a generalized jackknife estimate of bias analogous to $\widehat{B}_{\mathcal{T}_2,n,\kappa}$ obtained by replacing $\mathcal{T}_2$ with Υ in (4.9)–(4.12). Similar to (4.18), $\widehat{\Upsilon}_n$ can be viewed as an estimator of $\Upsilon = \exp(\mu_{\Xi})$. The results in the remainder of this sections are stated in terms of Ξ and Υ, with the understanding that these are place-holders as described above. The next result permits inference about $\Upsilon = \exp(\mu_{\Xi})$.

Theorem 4.10. *Under the conditions of Theorem 4.2, as $n \to \infty$*

$$n^{1/2}\left(\widehat{\Upsilon}_n - \widehat{B}_{\Upsilon,n,\kappa} - \exp(\mu_{\Xi}) + \xi_{\Upsilon,n,\kappa}\right) \xrightarrow{d} N\left(0,\ \exp(2\mu_{\Xi})\sigma_{\Xi}^2\right) \quad (4.37)$$

as $n \to \infty$ whenever $\kappa \geq 2/5$. In addition, for cases where $\kappa < 1/2$,

$$n^{\kappa}\left(\widehat{\Upsilon}_{n_{\kappa}} - \widehat{B}_{\Upsilon,n,\kappa} - \exp(\mu_{\Xi}) - \xi_{\Upsilon,n,\kappa}\right) \xrightarrow{d} N\left(0, \exp(2\mu_{\Xi})\sigma_{\Xi}^2\right) \quad (4.38)$$

as $n \to \infty$, where $\xi_{\Upsilon,n,\kappa} = O(n^{-\nu_1}(\log n)^{\nu_1}) = o(n^{-\kappa})$ and ν_1 is defined in Theorem 3.2.

For cases where $\kappa \geq 2/5$, Theorem 4.10 permits construction of an asymptotically correct $(1-\alpha)$ confidence interval for $\exp(\mu_{\Xi})$

Table 1
Productivity Change and Its Components for Swedish Pharmacies, 1980–1989 ($p = q = 4$).

Period	$\widehat{\mathcal{M}}_n$	$\widehat{\mathcal{E}}_{1,n}$	$\widehat{\mathcal{E}}_{2,n}$	$\widehat{\mathcal{T}}_{1,n}$	$\widehat{\mathcal{T}}_{2,n}$	$\widehat{\mathcal{S}}_{1,n}$	$\widehat{\mathcal{S}}_{2,n}$	$\widehat{\mathcal{S}}_{3,n}$
1980–1981	0.99553	1.01093	1.00411	0.98476***	0.99945	1.00339***	0.97980**	0.96640
1981–1982	1.03636	0.96712***	0.97440***	1.07159***	1.07646***	0.99626***	0.98252	0.97683
1982–1983	1.01220	1.02702***	1.02041***	0.98557***	0.98172**	1.00324**	0.99727	0.96616
1983–1984	0.97079	0.99373	0.99808	0.97692**	0.97535	0.99782	0.99065***	0.97401***
1984–1985	1.02153*	1.00194	1.00112*	1.01955***	1.01712**	1.00041	0.99224**	0.97827
1985–1986	1.00942*	0.99339	0.99289	1.01614***	1.01308***	1.00025	0.99335***	0.97811
1986–1987	1.03270	1.00074***	1.00522***	1.03194	1.02593	0.99777	0.99173***	0.97971
1987–1988	1.02152***	1.00663***	1.00455***	1.01479	1.00935	1.00103*	0.99331**	0.97124**
1988–1989	1.02532***	1.00278**	0.99984**	1.02248***	1.02876***	1.00147	0.98484	0.97381
1980–1989	1.09186**	1.00326***	1.00003***	1.08832**	1.10708***	1.00161**	0.97410***	0.93949

Note: Significant differences from 1 at levels .1, .05 or. 01 are indicated by one, two or three asterisks, respectively.

given by

$$\left[\widehat{\Upsilon}_n - \widehat{B}_{\Upsilon,n,\kappa} \pm z_{1-\frac{\alpha}{2}} \frac{\exp(\widehat{\mu}_{\Xi,n})\widehat{\sigma}_{\Xi,n}}{n^{1/2}}\right]. \tag{4.39}$$

Alternatively, whenever $\kappa < 1/2$, Theorem 4.10 can be used to construct the asymptotically correct $(1-\alpha)$ confidence interval

$$\left[\widehat{\Upsilon}_{n_\kappa} - \widehat{B}_{\Upsilon,n,\kappa} \pm z_{1-\frac{\alpha}{2}} \frac{\exp(\widehat{\mu}_{\Xi,n})\widehat{\sigma}_{\Xi,n}}{n^{\kappa}}\right]. \tag{4.40}$$

Analogous to the discussion above, one could also replace $\widehat{\Upsilon}_{n_\kappa}$ with $\widehat{\Upsilon}_n$ in (4.40), with the coverage of the resulting interval converging to 1 as $n \to \infty$.

5. An empirical illustration

Färe et al. (1992) examine productivity change among $n = 42$ Swedish pharmacies over 1980–1989. Their model specifies $p = 4$ inputs and $q = 4$ outputs.[6] As noted in Section 2, Färe et al. (1992) estimate a Malmquist index based on input-oriented distance measures, and decompose their index analogously to the decomposition in (2.9). For each pair of years 1980–1981, 1981–1982, ..., 1988–1989 Färe et al. (1992) report geometric means of their estimated Malmquist indices as well as for their estimates of input-oriented analogues of the components $\mathcal{E}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ and $\mathcal{T}_1(Z_i^1, Z_i^2 \mid \Psi^1, \Psi^2)$ defined in (2.9).

Table 1 shows geometric means of estimated hyperbolic Malmquist indices as well as of the various components defined in (2.9), (2.10), (2.11) and (2.13) obtained using the Färe et al. (1992) data on Swedish pharmacies. Since Färe et al. (1992) work in the input orientation, the geometric means reported in columns labeled $\widehat{\mathcal{E}}_{1,n}$, $\widehat{\mathcal{T}}_{1,n}$ and $\widehat{\mathcal{M}}_n$ in Table 1 are equal to square roots of the corresponding geometric means reported by Färe et al. (1992) in their Tables 1–3. One, two or three asterisks in Table 1 indicate statistical significance of differences from 1 at levels .1, .05 and .01 (respectively). In discussing their results, Färe et al. (1992, p. 96) remark that "According to our results, we have had on average productivity gains in seven periods and productivity losses in two periods." In Table 1, geometric means of estimated Malmquist indices for year-to-year periods range from 0.9708 to 1.0364, but are significantly different (at the.05 level) from 1 only for 1984–1985 and 1985–1986 (at.1) and 1987–1988 and 1988–1989 (at.01). However, looking at 1980–1989, there is evidence of considerable (about 9 percent, significant at.05) change in productivity.

In the periods where geometric means of estimated Malmquist indices are not significantly different from 1, some of the components of productivity change are significantly different from 1 although they work to offset each other resulting in no significant change in productivity. For example, the geometric mean of estimated Malmquist indices among the year-to-year periods is largest—1.03636—for 1981–1982. Both of the geometric means for efficiency change estimates are significant and less than 1, while both of the geometric means for change in technology are significant and greater than 1. Combined, these results suggest that while the technology shifted upward between 1981 and 1982, the pharmacies did not become more productive, and consequently became less technically efficient. In other words, the technology shifted, but the pharmacies did not. Instead, they were left behind. The value of 0.99626 for $\widehat{\mathcal{S}}_{1,n}$ during 1981–1982 indicates a significant decrease in scale efficiency, but the value is perhaps not economically significant since it is numerically close to 1.

For 1980–1989, the estimated value $\widehat{\mathcal{E}}_{2,n}$ for (geometric) mean efficiency change is numerically very close to 1, but significantly different from 1 at the.01 level. This is due in part to the fact that the estimated variance $\widehat{\sigma}_{\varepsilon_2,n}$ is rather small (0.00198). Although the estimate is significantly different from 1, it is perhaps not economically meaningful. By contrast, the value of the technology-change estimate $\widehat{\mathcal{T}}_{2,n}$ is equal to 1.10708, and significant at the.01 level. Together, these two estimates imply that the technology shifted upward between 1980 and 1989, and the pharmacies also shifted upward, keeping pace with the technology. The product of the scale estimates $\widehat{\mathcal{S}}_{1,n}$ and $\widehat{\mathcal{S}}_{2,n}$ is less than 1 (the estimate $\widehat{\mathcal{S}}_{3,n}$ is also less than 1). Thus while the results in Table 1 suggest that technology improved by about 10.7 percent, and efficiency was largely unchanged, the scale effects offset a small part of the improvement in technology resulting in an increase in average productivity of about 9.2 percent.

It is important to note that much more is known about the statistical properties of DEA estimators today than was known when Färe et al. (1992) published their paper. Today, we know the convergence rate of the VRS-DEA estimator and its conical hull (under VRS) is $n^{2/(p+q+1)}$. Moreover, with only $n = 42$ observations in each year and $p + q = 8$ dimensions, the well-known curse of dimensionality is problematic. The "effective parametric sample size" defined by Wilson (2018) is only 5. Moreover, the hyperbolic free-disposal hull efficiency estimator yields 40–42 observations with efficiency estimates equal to 1 in each year, providing another indication that the number of dimensions is too large for the available number of observations to obtain meaningful estimates.

Performing an eigensystem decomposition of the moment matrix $X'X$ of the 4 inputs as discussed by Wilson (2018) indicates the ratio of the largest eigenvalue to the sum of the eigenvalues is 95.4. The similar ratio for the 4 outputs is 91.3, and for the 3 outputs excluding deliveries to hospitals the corresponding ratio is 96.5. With only 42 observations in each year, the simulation results of Wilson (2018) strongly suggest that mean-square error of the estimates will be reduced using the dimension-reduction method described by Daraio and Simar (2007) and Wilson (2018).

[6] The inputs are (i) labor input for pharmacists; (ii) labor input for technical staff; (iii) building services; and (iv) equipment services. The outputs are (i) drug deliveries to hospitals; (ii) prescription drugs for outpatient care; (iii) medical appliances for the handicapped; and (iv) over the counter goods. See Färe et al. (1992) for further details. We are grateful to the authors for making the data available.

Table 2
Productivity Change and Its Components for Swedish Pharmacies, 1980–1989 ($p = 1,\ q = 2$).

Period	$\widehat{\mathcal{M}}_n$	$\widehat{\mathcal{E}}_{1,n}$	$\widehat{\mathcal{E}}_{2,n}$	$\widehat{\mathcal{T}}_{1,n}$	$\widehat{\mathcal{T}}_{2,n}$	$\widehat{\mathcal{F}}_{1,n}$	$\widehat{\mathcal{F}}_{2,n}$	$\widehat{\mathcal{F}}_{3,n}$
1980–1981	1.04115***	0.95615***	0.96654***	1.08889***	1.07917***	0.99461	0.97483*	0.96916**
1981–1982	1.03009**	0.84746***	0.91362***	1.21550***	1.13658***	0.96311***	0.95838***	0.96331
1982–1983	1.01365	1.17883***	1.08312***	0.85988***	0.93245***	1.04325***	0.89944***	0.89269***
1983–1984	0.98084*	0.94881***	1.00099*	1.03376***	0.98498	0.97359***	0.97029***	0.97443***
1984–1985	1.02794***	1.06846***	1.03847***	0.96208***	0.98618***	1.01434***	0.92795***	0.92367***
1985–1986	1.01819***	0.97724***	0.98190***	1.04190***	1.03259***	0.99763	0.95523***	0.94924***
1986–1987	1.04266***	1.01199**	1.02699***	1.03031***	1.01493***	0.99267***	0.94699***	0.94580***
1987–1988	1.04217***	0.98086***	0.99191***	1.06250***	1.04111***	0.99441***	0.94142***	0.93119***
1988–1989	1.03364***	1.05425***	1.01015***	0.98044***	1.03202***	1.02160***	0.91462***	0.91926***
1980–1989	1.24841***	0.99030	1.00457	1.26064***	1.25113***	0.99287***	0.97006***	0.96290

Note: Significant differences from 1 at levels .1, .05 or .01 are indicated by one, two or three asterisks, respectively.

Table 3
Tests for Differences Between "Small" and "Large" Pharmacies, 1980–1989 ($p = 1,\ q = 2$).

Period	$\widehat{\mathcal{M}}_n$	$\widehat{\mathcal{E}}_{1,n}$	$\widehat{\mathcal{E}}_{2,n}$	$\widehat{\mathcal{T}}_{1,n}$	$\widehat{\mathcal{T}}_{2,n}$	$\widehat{\mathcal{F}}_{1,n}$	$\widehat{\mathcal{F}}_{2,n}$	$\widehat{\mathcal{F}}_{3,n}$
1980–1981	−2.330**	5.118***	7.260***	−20.902***	−8.109***	−2.690***	1.150	1.055
1981–1982	−2.140**	15.485***	14.521***	−18.404***	−15.095***	2.064**	−0.040	−1.399
1982–1983	−0.872	10.758***	11.472***	−19.794***	−14.631***	0.788	−2.978***	−3.033***
1983–1984	−3.868***	10.131***	10.820***	−50.138***	−16.404***	−2.267**	8.202***	1.309
1984–1985	−0.341	−5.104***	−1.329	11.833***	0.135	−5.804***	10.177***	10.340***
1985–1986	−2.330**	9.300***	9.199***	−16.606***	−19.068***	−6.044***	8.891***	10.258***
1986–1987	−1.874*	−4.168***	12.903***	7.457***	−16.214***	−18.736***	9.561***	11.004***
1987–1988	−0.367	5.506***	8.139***	−8.917***	−16.266***	−6.563***	8.767***	8.908***
1988–1989	0.393	−9.836***	−3.398***	31.042***	0.321	−2.841***	6.211***	6.623***
1980–1989	−3.340***	3.258***	0.557	−8.948***	−5.250***	3.402***	4.561***	−5.463***

Note: Significant differences from 0 at levels .1, .05 or .01 are indicated by one, two or three asterisks, respectively.

Table 2 show results analogous to those in Table 1 but obtained working in the reduced space with only the first principle component of the 4 inputs (based on the moment matrix of the inputs) and the first output (i.e., hospital deliveries) and the first principle component of the remaining 3 outputs so that $p = 1$ and $q = 2$. Comparing the results in Table 2 with those in Table 1, it is evident that more estimates are significantly different from 1 when dimensionality is reduced than when working in the full-dimensional space. This is due to the fact that dimension reduction allows more precise (i.e., reduced mean-square error) estimates. Moreover, the results for productivity change and for technology change over 1980–1989 are considerably larger when dimension reduction is used. This again is to be expected, since with dimension reduction fewer individual distance estimates are equal to 1 than in the full-dimensional space. The overall conclusions drawn by Färe et al. (1992), i.e., that productivity change among the pharmacies is driven mainly by technology change, remain valid in view of the results in Table 2. In fact, the results are stronger and more dramatic than those originally reported by Färe et al. (1992).

It is straightforward to extend the results obtained above in Section 4 can be used to make test of differences between groups of producers, similar to the test for differences in mean efficiency developed by Kneip et al. (2016, Section 3.1.1). Suppose there are two independent groups (labeled "a" and "b", with n_a and n_b observations in each of two time periods with $n_a + n_b = n$. For $j \in \{a, b\}$ let $n_{j,\kappa} = min(\lfloor n_a^{2\kappa} \rfloor, n_j)$. Let $\widehat{\mathcal{T}}_{2,n_j}$ and $\widehat{\mathcal{T}}_{2,n_{j,\kappa}}$ and denote the geometric means for technical change defined in (4.18) and (4.21) for group $j \in \{a, b\}$. Similarly, let $\widehat{B}_{\mathcal{T}_{2,n_{j,\kappa}}}$ denote the generalized jackknife estimate of bias for group $j \in \{a, b\}$. Let $\mu_{\mathcal{T}_{2,j}}$ be the expectation corresponding to (4.2) for group $j \in \{a, b\}$. Theorem 4.7 can be used to establish the following result.

Theorem 5.1. *Under the conditions of Theorem 4.2,*

$$\frac{\left(\widehat{\mathcal{T}}_{2,n_a} - \widehat{\mathcal{T}}_{2,n_b}\right) - \left(\widehat{B}_{\mathcal{T}_{2,n_a,\kappa}} - \widehat{B}_{\mathcal{T}_{2,n_b,\kappa}}\right) - \left(\exp(\mu_{\mathcal{T}_{2,a}}) - \exp(\mu_{\mathcal{T}_{2,b}})\right)}{\left(\frac{\exp(2\widehat{\mu}_{\mathcal{T}_{2,n_a}})\widehat{\sigma}_{\mathcal{T}_{2,n_a}}}{n_a} + \frac{\exp(2\widehat{\mu}_{\mathcal{T}_{2,n_b}})\widehat{\sigma}_{\mathcal{T}_{2,n_b}}}{n_b}\right)^{1/2}} \xrightarrow{d} N(0,1) \tag{5.41}$$

as $n \to \infty$ *whenever* $\kappa \geq 2/5$. *In addition, for cases where* $\kappa < 1/2$,

$$\frac{\left(\widehat{\mathcal{T}}_{2,n_{a,\kappa}} - \widehat{\mathcal{T}}_{2,n_{b,\kappa}}\right) - \left(\widehat{B}_{\mathcal{T}_{2,n_{a,\kappa}}} - \widehat{B}_{\mathcal{T}_{2,n_{b,\kappa}}}\right)\left(\exp(\mu_{\mathcal{T}_{2,a}}) - \exp(\mu_{\mathcal{T}_{2,b}})\right)}{\left(\frac{\exp(2\widehat{\mu}_{\mathcal{T}_{2,n_{a,\kappa}}})\widehat{\sigma}_{\mathcal{T}_{2,n_a}}}{n_{a,\kappa}} + \frac{\exp(2\widehat{\mu}_{\mathcal{T}_{2,n_{b,\kappa}}})\widehat{\sigma}_{\mathcal{T}_{2,n_b}}}{n_{b,\kappa}}\right)^{1/2}} \xrightarrow{d} N(0,1) \tag{5.42}$$

as $n \to \infty$.

Under the null hypothesis that technical change is the same in both groups, $(\exp(\mu_{\mathcal{T}_{2,a}}) - \exp(\mu_{\mathcal{T}_{2,b}})) = 0$, and Theorem 5.1 can be used to test this hypothesis. Using similar reasoning, the results in Theorem 4.10 can be used to establish results for the other components of productivity change analogous to those in Theorem 5.1. In addition, Theorems 4.5 and 4.6 of Kneip et al. (2018) can be used to construct similar results for the Malmquist index itself.

Working in the reduced space with the pharmacy data, we have $p + q = 3$ and hence $\kappa = 1/2$. For each pair of years a and b, we sort the pharmacies according to the value of the single input in year a, and then compare productivity change and its components for the smallest 21 pharmacies versus the largest 21 pharmacies. Table 3 gives results of these tests.

Among the 80 test statistics reported in Table 3, all but 13 are significantly different from 0. All of the significant statistics for productivity change are negative, indicating that the larger pharmacies had greater productivity improvement (or smaller decrease in 1983–1984) than the smaller pharmacies. The results for technical change are similar, indicating that the larger pharmacies drove innovation more than the smaller pharmacies. On the other hand, statistics for the efficiency change measures as well as the scale measures are positive and significant in many cases, suggesting that the smaller pharmacies became relatively more efficiency than their larger counterparts in many cases.

6. Summary and conclusions

Indices arising from various decompositions of Malmquist indices are widely used to measure changes in technology efficiency, technology, scale efficiency and other factors and are often

estimated by nonparametric DEA estimators. Until now, no theoretical results justifying inference about the sources of productivity change measured by these indices have been available, nor have theoretical results permitting valid inference using geometric means of these indices been available. These deficiencies are remedied by the present paper. Results enabling inference via the subsampling methods of Simar and Wilson (2011) for individual producers are provided. In addition, new CLT results are established to enable inference about overall or average changes in terms of geometric means. Moreover, as shown in Section 5, it is easy to use these new CLT results to test hypotheses regarding differences in changes in efficiency, changes in technology, etc. between groups of firms along the lines of Kneip et al. (2016).

We focus in this paper on hyperbolic measures to avoid issues of existence and numerical difficulties. Simar, Vanhems, and Wilson (2012) extend results on hyperbolic VRS-DEA estimators to directional VRS-DEA estimators, and the similar arguments can be used to extend the results obtained above to directional measures and estimates. Of course, input and output-oriented measures are special cases of the directional measure, and so the results obtained above also extend trivially to the input and output orientations. Note also that Luenberger indices are simply additive versions of the multiplicative Malmquist indices. Hicks–Moorsteen indices involve ratios of both input-oriented and output-oriented distance measures.[7] Consequently, the results obtained here extend easily to make inference about these indices, too.

Appendix A. Technical details

A1. Additional assumptions

The two additional assumptions that appear in this section appear as Assumptions 3.1 and 3.2 in Kneip et al. (2018). The first assumption is needed to ensure that estimators of $\theta(x, y|\Psi^t)$ and $\gamma(x, y|\Psi^t)$ are well-defined. The second assumption ensures that the cross-efficiency estimators $\theta(Z_i^2 \mid \Psi^1)$ and $\theta(Z_i^1 \mid \Psi^2)$ as well as $\gamma(Z_i^2 \mid \Psi^1)$ and $\gamma(Z_i^1 \mid \Psi^2)$ are well-defined. Before stating the assumptions, some discussion is presented to establish notation used in the first assumption. See Kneip et al. (2018) for additional discussion.

Note that for a point $(x, y) \in \mathcal{D}^t$ the input-oriented efficiency $\theta(x, y \mid \mathcal{C}(\Psi^t))$ can be written as[8]

$$\theta(x, y \mid \mathcal{C}(\Psi^t)) = \min_{a>0} \left\{ \frac{\theta(x, ay \mid \Psi^t)}{a} \mid (\theta(x, ay \mid \Psi^t)x,\ ay) \in \Psi^t \right\}. \tag{A.1}$$

In addition, let $a_{min}^{x,y} \in \mathbb{R}_+$ denote the smallest $a > 0$ such that

$$\theta(x, y \mid \mathcal{C}(\Psi^t)) = \frac{\theta(x, a_{min}^{x,y} y \mid \Psi^t)}{a_{min}^{x,y}} = \min_{a>0} \left\{ \frac{\theta(x, ay \mid \Psi^t)}{a} \mid (\theta(x, ay \mid \Psi^t)x, ay) \in \Psi^t \right\}. \tag{A.2}$$

Necessarily, $a_{min}^{x,y} \in \mathbb{R}_+$ is uniquely defined if Ψ^t is strictly convex.

Recall that due to Assumptions 2.4–2.7, the support of any observable data in each period t is some subset $\mathcal{D}^t \subset \Psi^t$. In other words, $\mathcal{D}^t$ is the "observable part" of Ψ^t. The difference between $\mathcal{D}^t$ and Ψ^t does not play an important role in Kneip et al. (2008, 2015, 2016) since Assumption 2.5 requires (i) $(\theta(x, y \mid \Psi^t)x, y) \in \mathcal{D}^t$ for $(\theta(x, y \mid \Psi^t)x, y) \in \mathcal{D}^t$ and (ii) $f(\theta(x,\ y|\Psi^t)x,\ y) > 0$. Here, however, the difference between $\mathcal{D}^t$ and Ψ^t is problematic for dealing with $\theta(x, y \mid \mathcal{C}(\Psi^t))$. Furthermore, in order to ensure that Malmquist indices are well-defined, $\mathcal{D}^t$ and $\mathcal{D}^s$ must "fit together" for different periods t, s. Therefore, some additional conceptual work is necessary.

Let

$$\mathcal{D}_{norm}^t := \left\{ \left(\frac{x}{\|x\|}, \frac{y}{\|y\|} \right) \mid (x, y) \in \mathcal{D}^t \right\}. \tag{A.3}$$

If $p + q = 2$ then trivially $\mathcal{D}_{norm}^t = \{(1, 1)\}$. But when $p + q > 2$, $\mathcal{D}_{norm}^t$ will quantify the set of all possible "directions" of vectors x and y where it is possible to define a frontier. Note that for any $(\widetilde{x}, \widetilde{y})$ with $\|\widetilde{x}\| = 1$ and $\|\widetilde{y}\| = 1$ and $(\widetilde{x}, \widetilde{y}) \notin \mathcal{D}_{norm}^t$, we necessarily have $\{a\widetilde{x}, b\widetilde{y} \mid a, b > 0\} \cap \mathcal{D}^t = \emptyset$. This means that "in the direction" of $(\widetilde{x}, \widetilde{y})$ it is not possible to define any type of identifiable efficiency measure, since there is no information about an efficient frontier in such directions.[9]

Introduction of $\mathcal{D}_{norm}^t$ is of particular importance in a dynamic context where efficiencies in two different time periods t and s are to be compared. Frontiers may change and we may have different supports $\mathcal{D}^t$ and $\mathcal{D}^s$ in the two periods. However, it is necessary that $\mathcal{D}_{norm}^t = \mathcal{D}_{norm}^s$. Otherwise, there will be observations in one period for which distance to the other-period frontier cannot be defined. In this case Malmquist indices will be undefined with non-zero, non-negligible probability.

On the other hand, for any $(\frac{x}{\|x\|}, \frac{y}{\|y\|}) \in \mathcal{D}_{norm}^t$ there exists a unique ray defining the corresponding part of the conical hull frontier $\mathcal{C}^\partial(\Psi^t)$. This can easily be seen by letting $(\frac{x}{\|x\|}, \frac{y}{\|y\|}) \in \mathcal{D}_{norm}^t$. In addition, for $a > 0$, define

$$\widetilde{g}_x\left(a\frac{y}{\|y\|}\right) := \min_{b>0} \left\{ b\frac{x}{\|x\|} \mid \left(b\frac{x}{\|x\|}, a\frac{y}{\|y\|} \right) \in \Psi^t \right\}. \tag{A.4}$$

Then there exists some $\alpha_{min}^{x,y} > 0$ such that

$$\frac{\widetilde{g}_x(\alpha_{min}^{x,y}\frac{y}{\|y\|})}{\alpha_{min}^{x,y}} = \min_{a>0} \left\{ \frac{\widetilde{g}_x(a\frac{y}{\|y\|})}{a} \mid \left(g_x\left(a\frac{y}{\|y\|}\right)\frac{x}{\|x\|},\ a\frac{y}{\|y\|} \right) \in \Psi^t \right\} \tag{A.5}$$

where $\alpha_{min}^{x,y} \in \mathbb{R}_+$ is necessarily uniquely defined if Ψ^t is strictly convex.[10]

Assuming that only values a leading to well-defined frontier points are taken into account, for any $(x, y) \in \mathcal{D}^t$ we now have

$$\min_{a>0} \frac{\theta(x, ay)}{a} = \min_{a>0} \frac{\widetilde{g}_x(\|y\| a\frac{y}{\|y\|})}{\|x\| a} = \frac{\|y\|}{\|x\|} \min_{a>0} \frac{\widetilde{g}_x(\|y\| a\frac{y}{\|y\|})}{\|y\| a} = \frac{\|y\|}{\|x\|} \frac{\widetilde{g}_x(\alpha_{min}^{x,y}\frac{y}{\|y\|})}{\alpha_{min}^{x,y}}, \tag{A.6}$$

and $a_{min}^{x,y}$ defined in (A.2) satisfies $a_{min}^{x,y} = \frac{\alpha_{min}^{x,y}}{\|y\|}$.

Obviously, all we can hope to estimate is the version of (A.5) where Ψ^t is replaced by the observable part $\mathcal{D}^t \subset \Psi^t$. If $\alpha_{min}^{x,y} \in \mathbb{R}_+$ is such that $(g_x(\alpha_{min}^{x,y}\frac{y}{\|y\|})\frac{x}{\|x\|}, \alpha_{min}^{x,y}\frac{y}{\|y\|}) \notin \mathcal{D}^t$, then it is impossible to estimate $\theta(x, y \mid \mathcal{C}(\Psi^t))$ consistently. Minimizing (A.5) with respect to $\mathcal{D}^t$ instead of Ψ^t will then lead to a

[7] See Färe, Grosskopf, and Margaritis (2008) for discussions of the Luenberger and Hicks–Moorsteen indices.

[8] For any efficiency estimator $\theta(x, y|\Psi^t)$ considered in this section we will use the following conventions: if $(x, y) \notin \Psi^t$ with $(bx, y) \in \Psi^t$ for some $b > 1$ we set $\theta(x, y \mid \Psi^t) = b\theta(bx, y \mid \Psi^t)$. Otherwise, $\theta(x, y|\Psi^t) := 1$ (or $\widehat{\theta}(x, y \mid \Psi^t) := 1$) whenever the set of all possible values satisfying the defining inequalities is the empty set. Asymptotically, this has negligible effect.

[9] Under the strong disposability assumed in Assumption 2.3, the DEA and CDEA estimators of $\theta(x, y|\Psi^t)$ and $\theta(x, y \mid \mathcal{C}(\Psi^t))$ described above are well-defined and can be computed, but they do not estimate anything that does not depend entirely upon Assumption 2.3 or that can be identified from data when $(x, y) \notin \mathcal{D}_{norm}^t$.

[10] Note that $\widetilde{g}_x\left(a\frac{y}{\|y\|}\right)$ corresponds to the function $g_x\left(0, a\frac{y}{\|y\|}\right)$ defined in Kneip et al. (2008). The coordinate system introduced in Kneip et al. (2008) is not needed here, but is required in the proofs that follow in Appendix A.

"boundary solution" $\alpha^* \in \mathcal{D}^t$ which is "as close as possible" to $\alpha^{x,y}_{min} \in \mathbb{R}_+$. This can only be avoided by assuming that $\mathcal{D}^t$ is large enough such that (when minimizing (A.5) over $\mathcal{D}^t$ instead of Ψ^t) the solution $a^{x,y}_{min} \in \mathbb{R}_+$ is in the interior of $\mathcal{D}^t$ in the sense that $(g_x((\alpha^{x,y}_{min} - \delta)\frac{y}{\|y\|})\frac{x}{\|x\|}, (\alpha^{x,y}_{min} - \delta)\frac{y}{\|y\|}) \in \mathcal{D}^t$ as well as $(g_x((\alpha^{x,y}_{min} + \delta)\frac{y}{\|y\|})\frac{x}{\|x\|}, (\alpha^{x,y}_{min} + \delta)\frac{y}{\|y\|}) \in \mathcal{D}^t$. Since $\mathcal{D}^t$ is almost strictly convex by Assumption 2.7, $\alpha^{x,y}_{min} \in \mathbb{R}_+$ is necessarily unique, and $\frac{\widetilde{g}_x((\alpha^{x,y}_{min} - \delta)\frac{y}{\|y\|})}{(\alpha^{x,y}_{min} - \delta)} > \frac{\widetilde{g}_x(\alpha^{x,y}_{min}\frac{y}{\|y\|})}{\alpha^{x,y}_{min}}$ as well as $\frac{\widetilde{g}_x((\alpha^{x,y}_{min} + \delta)\frac{y}{\|y\|})}{(\alpha^{x,y}_{min} + \delta)} > \frac{\widetilde{g}_x(\alpha^{x,y}_{min}\frac{y}{\|y\|})}{\alpha^{x,y}_{min}}$. Convexity of Ψ^t then necessarily implies that this value $\alpha^{x,y}_{min} \in \mathbb{R}_+$ also corresponds to the solution of the original minimization problem with respect to Ψ^t. In this sense the following assumption ensures well-defined estimators of $\theta(x, y \mid \mathcal{C}(\Psi^t))$.

Assumption A.1. (i) The support $\mathcal{D}^t \subset \Psi^t$ of f is such that for any $(\frac{x}{\|x\|}, \frac{y}{\|y\|}) \in \mathcal{D}^t_{norm}$ we have $(\widetilde{g}_x(\alpha^{x,y}_{min}\frac{y}{\|y\|})\frac{x}{\|x\|}, \alpha^{x,y}_{min}\frac{y}{\|y\|}) \in \mathcal{D}^t$; (ii) there exists a $\delta > 0$ such that for any $(\frac{x}{\|x\|}, \frac{y}{\|y\|}) \in \mathcal{D}^t_{norm}$ we also have $(\widetilde{g}_x([\alpha^{x,y}_{min} - \delta]\frac{y}{\|y\|})\frac{x}{\|x\|}, [\alpha^{x,y}_{min} - \delta]\frac{y}{\|y\|}) \in \mathcal{D}^t$ and $(\widetilde{g}_x([\alpha^{x,y}_{min} + \delta]\frac{y}{\|y\|})\frac{x}{\|x\|}, [\alpha^{x,y}_{min} + \delta]\frac{y}{\|y\|}) \in \mathcal{D}^t$; (iii) There exists a constant $0 < M < \infty$ such that $\|x\| \leq M$ for all $(x, y) \in \mathcal{D}^t$.

Now turn to the dynamic case. Suppose that for two different time periods $t \in \{1, 2\}$ we have the set $\mathcal{X}_n = \{(Z^1_i), (Z^2_i)\}_{i=1}^n$ defined earlier in Section 2 of independent, identically distributed (iid) pairs (of pairs) of input and output quantities for the two different periods. In each period there may exist additional observations which do not possess a counterpart in the other period. More precisely, there are $n_1 \geq n$ observations in period 1 which are used to estimate the hyperbolic distance $\gamma^1(x, y) := \gamma(x, y \mid \mathcal{C}(\Psi^1))$, while there are $n_2 \geq n$ observations in period 2 which are used to estimate the hyperbolic distance $\gamma^2(x, y) := \gamma(x, y \mid \mathcal{C}(\Psi^2))$.

Assumption A.2. (i) For $t \in \{1, 2\}$ there are iid observations (X^t_i, Y^t_i), $i = 1, \ldots, n_t$, such that Assumptions 2.1–2.7 and A.1 are satisfied with respect to the underlying densities f^t with supports $\mathcal{D}^t$; (ii) $\mathcal{D}^1_{norm} = \mathcal{D}^2_{norm}$; (iii) for some $n \leq \min\{n_1, n_2\}$ the observations $\big((Z^1_i), (Z^2_i)\big)$, $i = 1, \ldots, n$ are iid and their joint distribution possesses a continuous density f_{12} with support $\mathcal{D}^1 \times \mathcal{D}^2$; (iv) for any $i = 1, \ldots, n$, (Z^1_i) is independent of (X^2_j, Y^2_j) for all $j = 1, \ldots, n_2$ with $i \neq j$; (v) for any $i = 1, \ldots, n$, (Z^2_i) is independent of (X^1_j, Y^1_j) for all $j = 1, \ldots, n_1$ with $i \neq j$.

Note that condition (i) of this assumption only guarantees that all estimators $\theta(x, y \mid \widehat{\Psi}^t_{n_t}))$ and $\gamma(x, y \mid \widehat{\Psi}^t_{n_t})$ follow the asymptotic distributions derived in Theorems 3.1 and 3.2 of Kneip et al. (2018). Condition (ii) together with Eq. (3.9) of Kneip et al. (2018) ensures that the cross-efficiency estimators $\theta(Z^2_i \mid \widehat{\Psi}^1_{n_1})$ and $\theta(Z^1_i \mid \widehat{\Psi}^2_{n_2})$ as well as $\gamma(Z^2_i \mid \widehat{\Psi}^1_{n_1})$ and $\gamma(Z^1_i \mid \widehat{\Psi}^2_{n_2})$ are asymptotically well-defined and possess the same rates of convergence as the contemporaneous efficiency estimators. Conditions (iv)–(v) permit dependence of a given firm's input-output quantities across periods 1 and 2, but require independence of the firm's input-output quantities from those of other firms in other periods.

A2. Proof of Theorem 3.1

For case (i) where $\Gamma(\cdot)$ denotes the identity function, the result follows immediately from Wilson (2011, Theorems 6.3.1 and 6.3.2). Given the result for case (i), the result for case (ii) where $\Gamma(\cdot)$ denotes the log function follows via the delta method given the fact that the log function is monotone and differentiable with non-zero derivatives on $\mathbb{R}_+$. □

A3. Proof of Theorem 3.2

For case (i) where $\Gamma(\cdot)$ denotes the identity function, consider the mapping ϕ from $\mathbb{R}^p_+ \times \mathbb{R}^q_+$ to $\mathbb{R}^p_+ \times \mathbb{R}^q_+$ such that $\phi : (x, y) \mapsto (x, y^{-1})$ where y^{-1} is the vector whose elements are the inverses of the corresponding elements of y. Denote $\omega = \phi(x, y)$. Clearly, ϕ is a continuous, one-to-one transformation; hence $(x, y) = \phi^{-1}(\omega)$. From the proof of Theorem 6.3.1 in Wilson (2011), it is clear that in ω-space, $\gamma(X_i, Y_i \mid \Psi^t)$ is an input-oriented efficiency measure along the lines of (2.4). Moreover, by Theorem 6.3.1 and Lemma 6.3.1 of Wilson (2011), $\gamma(X_i, Y_i \mid \widehat{\Psi}^t_n)$ is an ordinary input-oriented (VRS) DEA estimator along the lines of (3.2) with $(p + q)$ "inputs" and no outputs. Hence the results in (3.9)–(3.11) follow from Kneip et al. (2015, Theorem 3.1).

For case (ii) where $\Gamma(\cdot)$ denotes the log function, by A.1 part (iii), $\gamma(X_i, Y_i \mid \Psi^t)$ as well as the derivatives $\gamma'(X_i, Y_i \mid \Psi^t)$ and $\gamma''(X_i, Y_i \mid \Psi^t)$ are uniformly bounded for all $(X_i, Y_i) \in \mathcal{D}^t$. Then the results in (3.9)–(3.11) follow from arguments parallel to those used in the proof of Theorem 3.2 of Kneip et al. (2018). □

A4. Proof of Theorem 3.3

For $t = s$ the results in (3.12)–(3.14) follow trivially from case (ii) of Theorem 3.2.

For $t \neq s$, note that due to Assumption A.2, $\mathcal{D}^1_{\text{norm}} = \mathcal{D}^2_{\text{norm}}$. The results follow from arguments parallel to those used in the proof of Theorem 3.4 in Kneip et al. (2018). □

A5. Proof of Theorem 4.1

By definition, taking logs yields

$$\begin{aligned}&\log\left(\mathcal{T}_2(z^1, z^2 \mid \Psi^1, \Psi^2)\right)\\ &= \log\left(\gamma(z^2 \mid \Psi^1)\right) - \log\left(\gamma(z^2 \mid \Psi^2)\right)\\ &\quad + \log\left(\gamma(z^1 \mid \Psi^1)\right) - \log\left(\gamma(z^1 \mid \Psi^2)\right)\end{aligned} \tag{A.7}$$

and

$$\begin{aligned}&\log\left(\mathcal{T}_2(z^1, z^2 \mid \widehat{\Psi}^1_{n_1}, \widehat{\Psi}^2_{n_2})\right)\\ &= \log\left(\gamma(z^2 \mid \widehat{\Psi}^1_{n_1})\right) - \log\left(\gamma(z^2 \mid \widehat{\Psi}^2_{n_2})\right)\\ &\quad + \log\left(\gamma(z^1 \mid \widehat{\Psi}^1_{n_1})\right) - \log\left(\gamma(z^1 \mid \widehat{\Psi}^2_{n_2})\right).\end{aligned} \tag{A.8}$$

Note that Theorem 3.1 holds for both z^1 and z^2 due to Assumption A.2. Then

$$\begin{aligned}&n^\kappa\left(\log\left(\mathcal{T}_2(z^1, z^2 \mid \widehat{\Psi}^1_{n_1}, \widehat{\Psi}^2_{n_2})\right) - \log\left(\mathcal{T}_2(z^1, z^2 \mid \Psi^1, \Psi^2)\right)\right)\\ &\xrightarrow{\mathcal{L}} Q^{\log}_{\mathcal{T}_2, z^1, z^2}\end{aligned} \tag{A.9}$$

follows trivially from Theorem 3.1. The exponential function is monotonic and differentiable with nonzero derivatives on $\mathbb{R}_+$. Therefore the result follows from (A.9) via the delta method. □

A6. Proof of Theorem 4.2

First, let

$$\begin{aligned}\mathcal{R}_n &= E(\widehat{\mu}_{\mathcal{T}_2, n} - \mu_{\mathcal{T}_2})\\ &= D_{\mathcal{T}_2} n^{-\kappa} - \xi_{n,\kappa}.\end{aligned} \tag{A.10}$$

To simplify notation, let $\mathcal{T}_{2i} = \mathcal{T}_2(Z^1_i, Z^2_i \mid \Psi^1, \Psi^2)$ and let $\widehat{\mathcal{T}}_{2i} = \mathcal{T}_2(Z^1_i, Z^2_i \mid \widehat{\Psi}^1_{n_1}, \widehat{\Psi}^2_{n_2})$. Then (4.4) can be rewritten as

$$\begin{aligned}&n^{1/2}(\widehat{\mu}_{\mathcal{T}_2, n} - \mu_{\mathcal{T}_2} - \mathcal{R}_n)\\ &= \frac{n^{1/2}}{n}\sum_{i=1}^n\left(\log\widehat{\mathcal{T}}_{2i} - \log\mathcal{T}_{2i} - E\left(\log\widehat{\mathcal{T}}_{2i}\right) + \mu_{\mathcal{T}_2}\right)\end{aligned}$$

$$+ \frac{n^{1/2}}{n} \sum_{i=1}^{n} (\log \mathcal{T}_{2i} - \mu_{\mathcal{T}_2}). \tag{A.11}$$

Results (3.9)–(3.10) in Theorem 3.2 imply

$$\frac{n^{1/2}}{n} \sum_{i=1}^{n} (\log \widehat{\mathcal{T}}_{2i} - \log \mathcal{T}_{2i} - E(\log \widehat{\mathcal{T}}_{2i}) + \mu_{\mathcal{T}_2}) \xrightarrow{p} 0. \tag{A.12}$$

Hence result (4.4) follows from the Lindeberg-Levy CLT.

Second, the results in (4.5) follows directly from (A.9) in the proof of Theorem 4.1. In particular, $\widehat{\sigma}^2_{\mathcal{T}_2,n} = n^{-1} \sum_{i=1}^{n} (\log(\widehat{\mathcal{T}}_{2i}) - \widehat{\mu}_{\mathcal{T}_2,n})^2 \xrightarrow{p} E[(\log \widehat{\mathcal{T}}_{2i})^2] - \mu^2_{\mathcal{T}_2,n} = \text{VAR}(\log \mathcal{T}_{2i}) + [E(\log \mathcal{T}_{2i})]^2 - \mu^2_{\mathcal{T}_2} = \sigma^2_{\mathcal{T}_2}$ since $[E(\log \mathcal{T}_{2i})]^2 - \mu^2_{\mathcal{T}_2} = 0$. □

A7. Proof of Theorem 4.3

The result follows immediately from Theorem 4.2 since the remainder term is of order $o(n^{-\kappa})$ and hence $n^{\kappa} \xi_{\mathcal{T}_2} = n^{\kappa} o(n^{-\kappa}) = o(1)$. Since $\widehat{\mu}_{\mathcal{T}_2,n}$ in (4.4) has been replaced with $\widehat{\mu}_{\mathcal{T}_2,n_\kappa}$ in (4.8), the scale factor needed to stabilize the variance is n^{κ}. □

A8. Proof of Theorem 4.4

The results follow trivially after substituting the jackknife bias estimator into (4.4) and (4.8). When $(p+q) = 4$ then $\kappa = 2/5$, and the remainder term in (4.4) is $O(n^{-3\kappa/2})$ ignoring the $(\log n)$ term which does not affect the rate. Moreover, $n^{1/2} O(n^{-3\kappa/2}) = O(n^{-1/10})$, while in (4.8) $n^{\kappa} \xi_{n,\kappa} = O(n^{-1/5})$. □

A9. Proof of Theorem 4.5

The result follows using the delta method. Define

$$\mathcal{R}_n = E(\widehat{\mu}_{\mathcal{T}_2,n} - \mu_{\mathcal{T}_2}) = D_{\mathcal{T}_2} n^{-\kappa} + \xi_{n,\kappa} \tag{A.13}$$

where κ is the remainder term defined in (4.4) in Theorem 4.2. A Taylor expansion yields

$$\begin{aligned} &n^{1/2}(\exp(\widehat{\mu}_{\mathcal{T}_2,n}) - \exp(\mu_{\mathcal{T}_2} + \mathcal{R}_n)) \\ &\quad = \exp(\mu_{\mathcal{T}_2} + \mathcal{R}_n) n^{1/2} (\widehat{\mu}_{\mathcal{T}_2,n} - \mu_{\mathcal{T}_2} - \mathcal{R}_n) + O_p(n^{-1/2}). \end{aligned} \tag{A.14}$$

Since $\mathcal{R}_n = O(n^{-\kappa})$, the result follows from a further Taylor expansion of $\exp(\mu_{\mathcal{T}_2} + \mathcal{R}_n)$ and result (4.4) in Theorem 4.2. □

A10. Proof of Theorem 4.6

The exponential function is monotonic and differentiable with nonzero derivatives on $\mathbb{R}_+$. Therefore the result follows from Theorem 4.3 via the delta method. □

A11. Proof of Theorem 4.7

The results follow trivially after substituting the jackknife bias estimator into (4.19) and (4.22). □

References

Ahn, Y.-H., & Min, H. (2014). Evaluating the multi-period operating efficiency of international airports using data envelopment analysis and the Malmquist productivity index. *Journal of Air Transport Management, 39*, 12–22.

Alam, I. (2001). A nonparametric approach for assessing productivity dynamics of large U. S. banks. *Journal of Money, Credit, and Banking, 33*, 121–139.

Aly, H. Y., Grabowski, R. G., Pasurka, C., & Rangan, N. (1990). Technical, scale, and allocative efficiencies in U.S. banking: an empirical investigation. *Review of Economics and Statistics, 72*, 211–218.

Andries, A. M. (2011). The determinants of bank efficiency and productivity growth in the central and eastern European banking systems. *Eastern European Economics, 49*, 38–59.

Armagan, G., Ozden, A., & Bekcioglu, S. (2010). Efficiency and total factor productivity of crop production at NUTS1 level in Turkey: Malmquist index approach. *Quality and Quantity, 44*, 573–581.

Balk, B. (2001). Scale efficiency and productivity change. *Journal of Productivity Analysis, 15*, 153–183.

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

Bassem, B. S. (2014). Total factor productivity change of MENA microfinance institutions: A Malmquist productivity index approach. *Economic Modelling, 39*, 182–189.

de Castro Lobo, M. S., Ozcan, Y. A., da Silva, A. C. M., Lins, M. P. E., & Fiszman, R. (2010). Financing reform and productivity change in Brazilian teaching hospitals: Malmquist approach. *Central European Journal of Operations Research, 18*, 141–152.

Chang, S.-J., Hsiao, H.-C., Huang, L.-H., & Chang, H. (2011). Taiwan quality indicator project and hospital productivity growth. *Omega, 39*, 14–22.

Chowdhury, H., Wodchis, W., & Laporte, A. (2011). Efficiency and technological change in health care services in Ontario: An application of Malmquist productivity index with bootstrapping. *International Journal of Productivity and Performance Management, 60*, 721–745.

Daraio, C., & Simar, L. (2007). *Advanced robust and nonparametric methods in efficiency analysis*. New York: Springer Science+Business Media, LLC.

Daraio, C., Simar, L., & Wilson, P. W. (2018). Central limit theorems for conditional efficiency measures and tests of the 'separability condition' in non-parametric, two-stage models of production. *The Econometrics Journal, 21*, 170–191. doi:10.1111/ectj.12103.

Egilmez, G., & McAvoy, D. (2013). Benchmarking road safety of U.S. states: A DEA-based Malmquist productivity index approach. *Accident Analysis and Prevention, 53*, 55–64.

Färe, R. (1988). *Fundamentals of Production Theory*. Berlin: Springer-Verlag.

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

Färe, R., Grosskopf, S., Lindgren, B., & Roos, P. (1994a). Productivity developments in Swedish hospitals: A Malmquist output approach. In A. Charnes, W. W. Cooper, A. Y. Lewin, & L. M. Seiford (Eds.), *Data envelopment analysis: Theory, methodology and applications*. Dordrecht: Kluwer Academic Publishers.

Färe, R., Grosskopf, S., & Lovell, C. A. K. (1985). *The measurement of efficiency of production*. Boston: Kluwer-Nijhoff Publishing.

Färe, R., Grosskopf, S., & Margaritis, D. (2008). Productivity and efficiency: Malmquist and more. In H. Fried, C. A. K. Lovell, & S. Schmidt (Eds.), *The measurement of productive efficiency* (pp. 522–621). Oxford: Oxford University Press.

Färe, R., Grosskopf, S., & Margaritis, D. (2011). Malmquist productivity indexes and DEA. In W. W. Cooper, L. M. Seiford, & J. Zhu (Eds.), *Handbook on Data Envelopment Analysis* (pp. 127–149). Boston, MA: Springer US.

Färe, R., Grosskopf, S., & Norris, M. (1997). Productivity growth, technical progress, and efficiency change in industrialized countries: Reply. *American Economic Review, 87*, 1040–1043.

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

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

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

Gray, H. L., & Schucany, R. R. (1972). *The generalized jackknife statistic*. New York: Marcel Decker, Inc..

Johnson, A. L., & McGinnis, L. F. (2009). The hyperbolic-oriented efficiency measure as a remedy to infeasibility of super efficiency models. *The Journal of the Operational Research Society, 60*, 1511–1517.

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

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

Kneip, A., Simar, L., & Wilson, P. W. (2016). Testing hypotheses in nonparametric models of production. *Journal of Business and Economic Statistics, 34*, 435–456.

Kneip, A., Simar, L., & Wilson, P. W. (2018). Infernece in dynamic, nonparametric models of production: Central limit theorems for Malmquist indices. Discussion paper #2018/10, Institut de Statistique, Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Liu, S.-T. (2010). Measuring and categorizing technical efficiency and productivity change of commercial banks in taiwan. *Expert Systems with Applications, 37*, 2783–2789.

Lovell, C. A. K. (2003). The decomposition of Malmquist productivity indexes. *Journal of Productivity Analysis, 20*, 437–458.

Ng, Y. C. (2011). The productive efficiency of Chinese hospitals. *China Economic Review, 22*, 428–439.

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

Russell, R. R. (2018). Theoretical productivity indices. In E. Grifell-Tatjé, C. A. K. Lovell, & R. C. Sickles (Eds.), *The Oxford Handbook of Productivity Analysis* (pp. 101–138). New York: Oxford University Press.

Shephard, R. W. (1970). *Theory of cost and production functions*. Princeton: Princeton University Press.

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

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

Simar, L., & Wilson, P. W. (1999). Estimating and bootstrapping Malmquist indices. *European Journal of Operational Research, 115*, 459–471.

Simar, L., & Wilson, P. W. (2011). Inference by the *m* out of *n* bootstrap in nonparametric frontier models. *Journal of Productivity Analysis, 36*, 33–53.

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

Wilson, P. W. (2011). Asymptotic properties of some non-parametric hyperbolic efficiency estimators. In I. Van Keilegom, & P. W. Wilson (Eds.), *Exploring research frontiers in contemporary statistics and econometrics* (pp. 115–150). Berlin: Springer-Verlag.

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

Woo, C., Chung, Y., Chun, D., Seo, H., & Hong, S. (2015). The static and dynamic environmental efficiency of renewable energy: A Malmquist index analysis of OECD countries. *Renewable and Sustainable Energy Reviews, 47*, 367–376.

Wu, A.-H., Cao, Y.-Y., & Liu, B. (2014). Energy efficiency evaluation for regions in China: An application of DEA and Malmquist indices. *Energy Efficiency, 7*, 429–439.

Zofio, J. L. (2007). Malmquist productivity index decompositions: A unifying framework. *Applied Economics, 39*, 2371–2387.

Zofio, J. L., & Lovell, C. A. K. (2001). Graph efficiency and productivity measures: an application to US agriculture. *Applied Economics, 33*, 1433–1442.