European Journal of Operational Research 283 (2020) 737–747

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Interfaces with Other Disciplines

Metafrontier productivity indices: Questioning the common convexification strategy

Qianying Jin[a,b,*], Kristiaan Kerstens[c], Ignace Van de Woestyne[d]

[a] Business School, Hunan University, Changsha, Hunan, China
[b] IESEG School of Management, 3 rue de la Digue, Lille F-59000, France
[c] IESEG School of Management, CNRS-LEM (UMR 9221), University of Lille, 3 rue de la Digue, Lille F-59000, France
[d] KU Leuven, Research unit MEES, Warmoesberg 26, Brussel B-1000, Belgium

ARTICLE INFO

Article history:
Received 24 May 2019
Accepted 12 November 2019
Available online 21 November 2019

Keywords:
Data envelopment analysis
Metafrontier
Malmquist
Hicks–Moorsteen

ABSTRACT

While the construction of metafrontiers based on the union of underlying group frontiers normally yields a non-convex metaset, a large majority in the literature seems to assume that a convexification strategy leads to a reasonable convex approximation of this non-convex metafrontier. However, Kerstens, O'Donnell, and Van de Woestyne (2019) recently deliver new results on the union operator on technologies under a variety of assumptions and empirically illustrate that such a convexification strategy is doubtful. The purpose of this contribution is to verify to which extent such a convexification strategy is tenable when computing the Malmquist and Hicks–Moorsteen productivity indices with respect to a metafrontier. Furthermore, the differences between the Malmquist and Hicks–Moorsteen productivity indices are investigated at the metafrontier level. This existing methodology is empirically applied on a secondary data under a wide variety of assumptions: we explore balanced and unbalanced data as well as constant and variable returns to scale. Anticipating our key results, we provide statistical evidence on the potential bias arising from applying the convexification strategy for the metafrontier productivity indices.

1. Introduction

Different organisations across industries, regions and countries may face different production possibilities at a given point in time as well as over time. Heterogeneity in performance can be due to differences in available technologies (i.e., the ways inputs can be transformed into outputs) and/or to differences in environments (e.g., economic infrastructure, regulation, geography, climate, etc.). There have been a variety of alternative proposals around to account for heterogeneity in production models. Some rather popular methods include the use of latent class models (e.g., Orea & Kumbhakar, 2004), the aggregation over groups or industries (e.g., Mayer & Zelenyuk, 2014, but see Balk, 2016 for some caveats), among others. To the best of our knowledge, no theoretical or empirical review has ever compared these different methods to account for heterogeneity in production.

This contribution focuses on one particular method to account for heterogeneity when estimating production relations. One historically important literature was initiated by Hayami and Ruttan (1970) who proposed and estimated a kind of meta-production function. This meta-production function concept has been empirically applied mainly in agriculture and for country-level data: an empirical survey is found in Trueblood (1989). Hayami and Ruttan (1970, p. 898) "call the envelope of all known and potentially discoverable activities a secular or "meta-production function"." This secular production function indicates the maximum output obtainable from given inputs and from a given stock of knowledge. Thus, all organisations have access to the same set of input-output combinations, but each may choose a different input-output combination from that set depending on specific circumstances (e.g., regulation, relative prices, etc.). Some of this literature takes the possibility of inefficiency into account (e.g., Lau & Yotopoulos, 1989).

These basic ideas have initially been transposed into a stochastic production frontier framework by Battese and Rao (2002) and Battese, Rao, and O'Donnell (2004). Thereafter, O'Donnell, Rao, and Battese (2008) refined the loose ends in the methodology and finalised the formal metafrontier framework for making efficiency comparisons across groups of firms using both stochastic frontier

* Corresponding author at: Business School, Hunan University, Changsha, Hunan, China.
E-mail addresses: qianyingjin@hnu.edu.cn, q.jin@ieseg.fr (Q. Jin), k.kerstens@ieseg.fr (K. Kerstens), ignace.vandewoestyne@kuleuven.be (I. Van de Woestyne).

https://doi.org/10.1016/j.ejor.2019.11.019
0377-2217/

analysis and nonparametric deterministic frontiers. This seminal article defines a meta-production possibility set (or metaset) as the union of underlying group-specific sets. These authors refer to the boundary of the metaset as a metafrontier, and to the boundaries of the group-specific sets as group-specific frontiers (or group frontiers).

This so-called metafrontier approach has meanwhile been amply applied across sectors and disciplines. Examples are production studies from agriculture (e.g., Latruffe, Fogarasi, & Desjeux, 2012), banking (e.g., Casu, Ferrari, & Zhao, 2013), fisheries (e.g., Lee & Midani, 2015), hotels (e.g., Huang, Ting, Lin, & Lin, 2013), schools (e.g., Thieme, Prior, & Tortosa-Ausina, 2013), and wastewater treatment plants (e.g., Sala-Garrido, Molinos-Senante, & Hernández-Sancho, 2011) to name but a few. This basic metafrontier concept has found its way in a variety of other literatures: one example is its transposition to a cost frontier framework (e.g., Huang & Fu, 2013); another example is the computation of productivity indices relative to metafrontiers (see, e.g., Casu et al., 2013 and Huang, Juo, & Fu, 2015 for a primal respectively a dual Malmquist index); a final example is the development of more elaborate efficiency decompositions (see Kounetas, Mourtos, & Tsekouras, 2009 and Tsekouras, Chatzistamoulou, & Kounetas, 2017).

Basic group-specific frontier models tend to make a series of standard assumptions, one of which is convexity. This convexity assumption can only be justified by a time divisibility argument (see Hackman, 2008, p. 39). But, even if group-specific sets are convex, then the metaset defined by their union is normally nonconvex (see O'Donnell et al., 2008). Despite this basic mathematical fact that convex group-specific sets yield a nonconvex metaset, the seminal article of O'Donnell et al. (2008) adopts a convexification strategy by estimating the metafrontier as a boundary of a convex metaset (see also, e.g., Battese & Rao, 2002; Battese et al., 2004). Since this convexification strategy is normally not true, estimates of the metafrontier are potentially biased. While the large majority of articles adopting a metafrontier approach seem to follow such a convexification strategy, one should stress that some articles do not adopt such a strategy: examples include Huang et al. (2013), Sala-Garrido et al. (2011), Tiedemann, Francksen, and Latacz-Lohmann (2011), and (partially) Walheer (2018) among others. Kerstens, O'Donnell, and Van de Woestyne (2019) elaborate on the union operation on technologies under various assumptions and find empirically convincing evidence that a convexification strategy leads to erroneous results in the estimation of efficiency measures.

The first purpose of this contribution is to investigate the impact of a convexification strategy on the estimation of metafrontier-based productivity indices. A variety of productivity indices have been computed using metafrontiers: examples include the popular primal Malmquist index (e.g., Casu et al., 2013) as well as the primal Luenberger indicator (e.g., Zhang & Wang, 2015), the dual (most often cost-based) Malmquist index (e.g., Huang et al., 2015), the primal Färe–Primont (e.g., Dakpo, Desjeux, Jeanneaux, & Latruffe, 2019), Hicks–Moorsteen (e.g., Verschelde, Dumont, Rayp, & Merlevede, 2016) and Lowe (e.g., O'Donnell, Fallah-Fini, & Triantis, 2017) Total Factor Productivity (TFP) indices, among others.[1] The second purpose is to contrast the differences between the Malmquist and Hicks–Moorsteen productivity indices at the metafrontier level rather than at the standard frontier level (see Kerstens & Van de Woestyne, 2014 for the latter comparison).

This contribution is structured as follows. The next Section 2 develops the geometric intuition why a convexification strategy may potentially lead to biases in the estimation of the metafrontier. Section 3 introduces notation and formally outlines the metafrontier methodology. Thereafter, Section 4 defines the productivity indices that are computed relative to the metafrontier: on the one hand the input-oriented Malmquist productivity index, and on the other hand the Hicks–Moorsteen TFP index that combines input- and output-oriented efficiency measures. Section 5 specifies the details about the deterministic nonparametric frontier technologies employed in computing the productivity indices relative to the metafrontier. Then, the next Section 6 offers an empirical illustration using a secondary data set of hydroelectric power plants from Chile. Finally, the last Section 7 summarizes results and draws some conclusions.

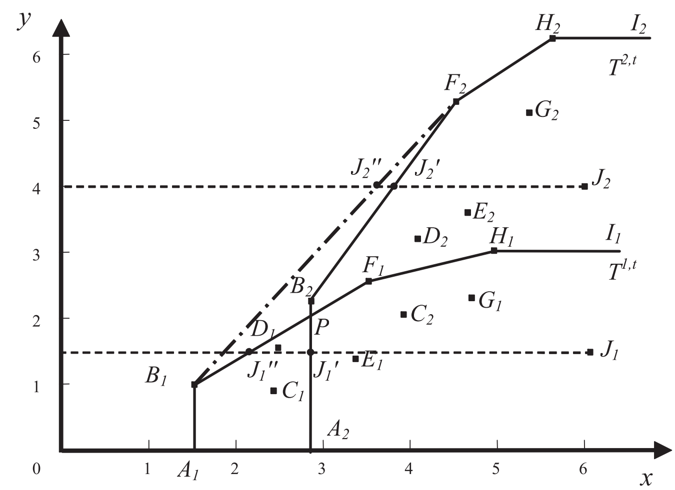

Fig. 1. Group technologies and metatechnology: the single-input–single-output case.

2. Metafrontier and convexification strategy: a clarification

It is essential to remind the reader about the intuition underlying the metafrontier approach. To fix our ideas, we start by an example taken from wastewater treatment plants (Sala-Garrido et al., 2011). There are four main technologies for wastewater treatment: activated sludge, aerated lagoon, trickling filter, and rotating biological contactor (biodisk). Suppose we focus on two such technologies to simplify matters.

Fig. 1 illustrates the single-input-single-output case when only two technologies exist. Group technology $T^{1,t}$ consists of 8 observations (B_1, C_1, D_1, E_1, F_1, G_1, H_1, J_1) denoted by square dots and represented by the polyline $A_1B_1F_1H_1I_1$ and the horizontal axis. Group technology $T^{2,t}$ consists of 8 observations (B_2, C_2, D_2, E_2, F_2, G_2, H_2, J_2) also denoted by square dots and represented by the polyline $A_2B_2F_2H_2I_2$ and the horizontal axis. The metatechnology $T^{\Gamma,t}$ is now the union of technologies $T^{1,t}$ and $T^{2,t}$: it is clearly nonconvex. $T^{\Gamma,t}$ consists of all points between the polyline $A_1B_1PB_2F_2H_2I_2$ and the horizontal axis. The convexification strategy consists in convexifying this nonconvex metatechnology $T^{\Gamma,t}$ by adding the region denoted by the polyline $B_1PB_2F_2B_1$.

Let us now explore what happens when we project inefficient observations with respect to the frontiers of these technologies. From the 8 observations in group technology $T^{1,t}$, 5 observations are situated below the frontier and are therefore inefficient. Let us focus on inefficient observation J_1: it is projected onto projection point J_1'' on the line segment B_1F_1 of the group technology $T^{1,t}$ and can learn to position itself onto this frontier from combining somehow the inputs and outputs of observations B_1 and F_1. The same observation J_1 can also be projected with respect to the other group technology at projection point J_1'.

From the 8 observations in group technology $T^{2,t}$, 5 observations are also situated below the frontier and are therefore inefficient.

[1] Recent surveys on productivity indices and indicators are found in O'Donnell (2018) and Russell (2018), among others.

Let us focus on inefficient observation J_2: it is projected onto projection point J_2' on the line segment B_2F_2 of the group technology $T^{2,t}$ and can learn to position itself onto this frontier from combining somehow the inputs and outputs of observations B_2 and F_2.

However, the same observation J_2 can now be projected with respect to the other group technology at projection point J_2'' depending on whether or not we adopt a convexification strategy. If we do not adopt a convexifixation strategy, then the metatechnology is nonconvex and the projection point J_2'' is simply infeasible. The distance to the metatechnology coincides with the distance to its own group technology and the distance to the other group technology is simply undefined. The reason why the projection point J_2'' is deemed infeasible is because this presupposes making a linear combination of point B_1 from group technology $T^{1,t}$ and point F_2 from group technology $T^{2,t}$. While we allow for convex combinations within each group technology, we normally rule out taking convex combinations across group technologies.

If we adopt a convexifixation strategy, then the metatechnology becomes convex again and the projection point J_2'' can be achieved. It should be realized that this convexification strategy is to some extent self-contradictory, because it destroys the very idea of distinguishing between different group technologies and only allowing for convexity per group technology. In other words, the union operator on group technologies does not normally preserve the convexity axiom on the resulting metatechnology.

The large majority of articles adopting a metafrontier approach seem to follow such a convexification strategy: a benign interpretation is that most authors just follow O'Donnell et al. (2008) and assume that such a strategy is rather harmless. Kerstens et al. (2019) cite a handful of articles that do not adopt such a strategy. Walheer (2018, p. 1015) states in this context: "All in all, the safest option is to assume a non-convex envelopment, as in the original definition of O'Donnell et al. (2008), while assuming a convex envelopment should be well-motivated."

Our reading of the metafrontier productivity literature also confirms this tendency: the large majority of articles adopts a convexification strategy. We are only aware of three exceptions. First, Verschelde et al. (2016) compute a metafrontier Hicks–Moorsteen TFP index starting from nonconvex group technologies: this choice for nonconvex technologies automatically leads to a nonconvex metatechnology. However, most researchers seem reluctant to give up the convexity of the group technologies and therefore we also maintain the convexity of the group technologies in this contribution.

Second, Afsharian, Ahn, and Harms (2018) report a metafrontier Malmquist productivity index starting from convex group technologies and report differences when computing a wrong convexified metatechnology rather than a correct nonconvex metatechnology. However, our work differs from these authors in two respects. First, we use formal statistical test procedures to verify whether a convexification strategy is innocuous or not (instead of a mere comparison). Second, we use the standard Malmquist productivity index computed over a two year time window, while these authors compute a so-called overall Malmquist index relative to one global technology computed over all available time periods (see Afsharian & Ahn, 2015 for more details).

Third, Walheer (2018) focuses on the aggregation of metafrontier technology gap ratios and contrasts the results of wrong convexified and correct nonconvex metatechnologies, but this author reports no formal statistical test procedure.

3. Metafrontier methodology

In this methodological section, we follow closely the notation and terminology introduced in Kerstens et al. (2019). We mainly introduce an additional time superscript to handle the time dynamics of productivity measurement.

3.1. Technology and technology-specific frontier, metatechnology and metafrontier

O'Donnell (2016, p. 328) defines a *technology* as "a technique, method or system for transforming inputs into outputs…For most practical intents and purposes, it is convenient to think of a technology as a book of instructions, or recipe". This definition is adopted here: we perceive a technology as a kind of intellectual capital.

Technology is represented by a *technology-specific production possibilities set* (TPPS), which is a set containing all possible combinations of inputs and outputs using a given technology. Let $x^t \in \mathbb{R}_+^M$ denote vectors of inputs and let $y^t \in \mathbb{R}_+^N$ denote vectors and outputs at time period t. The set of all pairs of input and output vectors that can be produced at time period t using technology g is described as follows:

$$T^{g,t}=\{(x^t, y^t) \in \mathbb{R}_+^M \times \mathbb{R}_+^N : x^t \text{with technology } g \text{ can produce } y^t\}. \tag{1}$$

The boundary of this TPPS is called a *technology-specific frontier*. Commonly, one makes the following assumptions on the TPPS:

(T.1) $(x^t, 0) \in T^{g,t}$ for all $x^t \in \mathbb{R}_+^M$.
(T.2) If $(0, y^t) \in T^{g,t}$, then $y^t = 0$.
(T.3) $T^{g,t}$ is a closed subset of $\mathbb{R}_+^M \times \mathbb{R}_+^N$.
(T.4) If $(x^t, y^t) \in T^{g,t}$ and $(x', -y') \geq (x^t, -y^t)$, then $(x', y') \in T^{g,t}$.
(T.5) $T^{g,t}$ is a convex set.
(T.6) If $(x^t, y^t) \in T^{g,t}$, then $\delta(x^t, y^t) \in T^{g,t}$ for all $\delta \geq 0$.

These traditional axioms concerning technology g in period t state that: (i) inaction is possible, (ii) there is no free lunch, (iii) the set of feasible input-output combinations contains all the points on its boundary (closedness), (iv) inputs and outputs are strongly (or freely) disposable, (v) the technology is convex, and (vi) the technology satisfies constant returns to scale in that observations can be scaled down or up at will. For more details on these axioms: see, for instance, Hackman (2008).

Note that the first assumption (T.1) is not always maintained in this contribution. Furthermore, in the empirical illustration we impose either constant returns to scale (T.6) or the more traditional variable returns to scale assumption (which amounts to the absence of any scaling: $\delta = 1$).

When axiom (T.4) is maintained, then $T^{g,t}$ is represented by the following technology-specific input distance function:

$$d_I^{g,t}(x^t, y^t) = \sup_{\lambda \in \mathbb{R}_+} \left\{\lambda : (x^t/\lambda, y^t) \in T^{g,t}\right\}. \tag{2}$$

This function is (i) non-negative, (ii) linearly homogeneous in inputs, and (iii) no less than unity for all $(x^t, y^t) \in T^{g,t}$.[2]

The *technology set* or *metatechnology* Γ is the set of all technologies g that exist for all time periods. If a technology is seen as a recipe, then following Caselli and Coleman (2006, p. 509) one can view a technology set as "a library, containing blueprints, or recipes to turn inputs into outputs". The set of all input and output vectors that are feasible using a given technology set Γ (i.e., using some technology that is contained in Γ) is labelled a *metatechnology-specific production possibilities set* (MTPPS). Mathematically, this MTPPS is defined as

$$T^{\Gamma,t} = \{(x^t, y^t) \in \mathbb{R}_+^M \times \mathbb{R}_+^N : \exists g \in \Gamma \text{ such that } x^t \text{ with technology } g \text{ can produce } y^t\}. \tag{3}$$

[2] Färe and Primont (1995, p. 22) indicate that weak rather than strong disposability of the inputs is sufficient to guarantee this representation.

Obviously, we have that $T^{\Gamma,t} = \cup_{g\in\Gamma} T^{g,t}$. The boundary of a MTPPS is called a *metafrontier*.

When strong disposability applies (i.e., (T.4) is true), then the MTPPS $T^{\Gamma,t}$ can be represented using the metatechnology-specific input distance function:

$$D_I^{\Gamma,t}(x^t, y^t) = \max_{g\in\Gamma}\{d_I^{g,t}(x^t, y^t)\}. \tag{4}$$

Equivalently, $D_I^{\Gamma,t}(x^t, y^t) = \sup_{\lambda\in\mathbb{R}_+}\{\lambda : (x^t/\lambda, y^t) \in T^{\Gamma,t}\}$. This function is non-negative, linearly homogeneous in inputs, and no less than unity for all $(x^t, y^t) \in T^{\Gamma,t}$.

Instead of using the technology-specific input distance function (2), $T^{g,t}$ can also be represented by the technology-specific output distance:

$$d_O^{g,t}(x^t, y^t) = \inf_{\lambda\in\mathbb{R}_+}\left\{\lambda : (x^t, y^t/\lambda) \in T^{g,t}\right\}. \tag{5}$$

This function is (i) non-negative, (ii) linearly homogeneous in outputs, and (iii) no greater than unity for all $(x^t, y^t) \in T^{g,t}$.[3] Under the strong disposability assumption (T.4), the MTPPS $T^{\Gamma,t}$ can then be represented using the metatechnology-specific output distance function:

$$D_O^{\Gamma,t}(x^t, y^t) = \min_{g\in\Gamma}\{d_O^{g,t}(x^t, y^t)\}. \tag{6}$$

Equivalently, $D_O^{\Gamma,t}(x^t, y^t) = \inf_{\lambda\in\mathbb{R}_+}\{\lambda : (x^t, y^t/\lambda) \in T^{\Gamma,t}\}$. This function is non-negative, linearly homogeneous in outputs, and less than unity for all $(x^t, y^t) \in T^{\Gamma,t}$.

3.2. Technical efficiency

In this contribution, the input-oriented metatechnology-specific technical efficiency (*ITE*) of an organization using inputs x^t to produce outputs y^t using some technology $g \in \Gamma$ at time period t is defined as the reciprocal of the metatechnology-specific input distance function (4):

$$\text{ITE}^{\Gamma,t}(x^t, y^t) = 1/D_I^{\Gamma,t}(x^t, y^t). \tag{7}$$

This radial technical efficiency measure lies in the closed unit interval and indicates the maximum proportional reduction in x^t that still allows production of y^t by some technology $g \in \Gamma$.

If Γ contains more than one technology, then the measure of *ITE* (7) can be written as the product of an input-oriented metatechnology ratio (*IMR*) and a measure of residual input-oriented technical efficiency (*RITE*). Mathematically, the *IMR* relative to the technology set Γ of a firm that uses inputs x^t and technology g to produce outputs y^t is

$$\text{IMR}^{g\Gamma,t}(x^t, y^t) = d_I^{g,t}(x^t, y^t)/D_I^{\Gamma,t}(x^t, y^t). \tag{8}$$

Also this measure lies in the closed unit interval. It can be interpreted as an input-oriented technical efficiency measure of whether a firm has chosen the best technology that is available. The associated measure of *RITE* is

$$\text{RITE}^{g,t}(x^t, y^t) = 1/d_I^{g,t}(x^t, y^t). \tag{9}$$

This measure also lies in the closed unit interval and indicates the maximum proportional reduction in x^t that allows production of y^t when using technology g for time period t. It can also be interpreted as the component of *ITE* that remains after accounting for the *IMR* (whence the term "residual"). Obviously, Eqs. (4), (7) and (9) imply that

$$\text{ITE}^{\Gamma,t}(x^t, y^t) = \min_{g\in\Gamma}\{\text{RITE}^{g,t}(x^t, y^t)\}. \tag{10}$$

[3] Färe and Primont (1995, p. 22) show that weak rather than strong disposability of the outputs is sufficient to guarantee this representation.

Note that some of the components in (10) can be undefined for some input-output combinations that are not contained in the group technology composing the technology or metatechnology (see Briec & Kerstens, 2009 for details on infeasibilities). Finally, Eqs. (7)–(9) imply that

$$\text{ITE}^{\Gamma,t}(x^t, y^t) = \text{IMR}^{g\Gamma,t}(x^t, y^t) \cdot \text{RITE}^{g,t}(x^t, y^t). \tag{11}$$

Hence, technical efficiency can be decomposed into the product of a metatechnology ratio and a measure of residual technical efficiency: the first measures how close a technology-specific frontier is to the metafrontier, while the second measures how close a firm is operating to the technology-specific frontier.

By analogy with the former, the output-oriented metatechnology-specific technical efficiency (*OTE*) of an organization using inputs x^t to produce outputs y^t using some technology $g \in \Gamma$ at time period t is defined as the reciprocal of the metatechnology-specific output distance function (6):

$$\text{OTE}^{\Gamma,t}(x^t, y^t) = 1/D_O^{\Gamma,t}(x^t, y^t). \tag{12}$$

This radial technical efficiency measure results in values larger than or equal to one and indicates the maximum proportional expansion in y^t that is still achievable with x^t inputs by some technology $g \in \Gamma$.

If Γ contains more than one technology, then the measure of *OTE* (12) can be written as the product of an output-oriented metatechnology ratio (*OMR*) and a measure of residual output-oriented technical efficiency (*ROTE*). Mathematically, the *OMR* relative to the technology set Γ of a firm that uses inputs x^t and technology g to produce outputs y^t is

$$\text{OMR}^{g\Gamma,t}(x^t, y^t) = d_O^{g,t}(x^t, y^t)/D_O^{\Gamma,t}(x^t, y^t). \tag{13}$$

The associated measure of *ROTE* is

$$\text{ROTE}^{g,t}(x^t, y^t) = 1/d_O^{g,t}(x^t, y^t). \tag{14}$$

This measure also has values greater than or equal to one and indicates the maximum proportional expansion in y^t that can still be realized with inputs x^t when using technology g for time period t. It can also be interpreted as the component of *OTE* that remains after accounting for the *OMR* (whence the term "residual"). Obviously, Eqs. (6), (12) and (14) imply that

$$\text{OTE}^{\Gamma,t}(x^t, y^t) = \max_{g\in\Gamma}\{\text{ROTE}^{g,t}(x^t, y^t)\}. \tag{15}$$

Like in the input orientation, some of the components in (15) can be undefined for some input-output combinations that are not contained in the group technology composing the technology or metatechnology (see Briec & Kerstens, 2009 for details on infeasibilities).

4. Metafrontier productivity indices

The measurement of productivity has in the last 25 years or so often been analysed using a technology-based, discrete-time Malmquist productivity index. Initially defined by Caves, Christensen, and Diewert (1982) as a ratio of distance functions, this index has become increasingly popular due to the innovations of Färe, Grosskopf, Lindgren, and Roos (1995). The latter authors have shown how: (i) to relax the hypothesis of technical efficiency maintained in Caves et al. (1982); (ii) to decompose this index into technology shifts and technical efficiency changes; and (iii) to compute this index relative to multiple inputs and outputs technologies by exploiting the relationship between distance functions and technical efficiency measures. O'Donnell (2012) argues rather convincingly that the Malmquist productivity index is not a Total Factor Productivity (TFP) index. This same position is also found in, among others, O'Donnell (2018, p. 120–121) and Russell (2018). It is rather a technology index aimed at mainly measuring local technical change (see Grosskopf, 2003). Therefore, we also compare the Malmquist productivity index with the Hicks–Moorsteen TFP index, one among the several available TFP indices (see O'Donnell, 2018; Russell, 2018).

To the best of our knowledge, this is the first comparison of the Malmquist productivity index and the Hicks–Moorsteen TFP index within a metafrontier context. Earlier, such a comparison using standard technologies has already been published in the literature (see e.g., Kerstens & Van de Woestyne, 2014). The latter authors report that the Malmquist and Hicks–Moorsteen indices are empirically clearly distinct under variable returns to scale. Under constant returns to scale, empirical differences are less clear cut at the sample level, though strong differences may persist for individual observations.[4]

4.1. Metafrontier Malmquist productivity index

One can define the input-oriented metafrontier Malmquist productivity index (*IMMI*) in base period t as follows:

$$IMMI^{\Gamma,t}(x^t,y^t,x^{t+1},y^{t+1}) = \frac{ITE^{\Gamma,t}(x^{t+1},y^{t+1})}{ITE^{\Gamma,t}(x^t,y^t)}. \tag{16}$$

Values of this base period t input-oriented *IMMI* above (below) unity reveal productivity growth (decline). Similarly, a base period $t+1$ input-oriented *IMMI* is defined as follows:

$$IMMI^{\Gamma,t+1}(x^t,y^t,x^{t+1},y^{t+1}) = \frac{ITE^{\Gamma,t+1}(x^{t+1},y^{t+1})}{ITE^{\Gamma,t+1}(x^t,y^t)}. \tag{17}$$

Again, values of this base period $t+1$ input-oriented *IMMI* above (below) unity reveal productivity growth (decline).

To avoid an arbitrary selection among base years, the input-oriented *IMMI* is defined as a geometric mean of a period t and a period $t+1$ index:

$$IMMI^{\Gamma,t,t+1}(x^t,y^t,x^{t+1},y^{t+1}) = \left[\frac{ITE^{\Gamma,t}(x^{t+1},y^{t+1})}{ITE^{\Gamma,t}(x^t,y^t)} \cdot \frac{ITE^{\Gamma,t+1}(x^{t+1},y^{t+1})}{ITE^{\Gamma,t+1}(x^t,y^t)}\right]^{1/2}. \tag{18}$$

Note again that when the geometric mean of the *IMMI* is larger (smaller) than unity, then it points to a productivity growth (decline).

Remark that the above definitions deviate from the original ones in Caves et al. (1982) in that the ratios have been inverted. This ensures that productivity indices above (below) unity reveal productivity growth (decline), which is in line with traditional TFP indices.

There is a considerable literature and quite some controversy on the best way to decompose the Malmquist productivity index (see, e.g., Lovell, 2003). In this contribution, we opt for the simplest possible decomposition of the metafrontier Malmquist productivity index and we refrain from entering into these controversies on the best decomposition (see Zofio, 2007 for a somewhat dated summary of these discussions). Following Färe, Grosskopf, Lindgren, and Roos (1992) and Färe, Grosskopf, Norris, and Zhang (1994), an equivalent way of writing this *IMMI* index is

$$IMMI^{\Gamma,t,t+1}(x^t,y^t,x^{t+1},y^{t+1}) = \underbrace{\frac{ITE^{\Gamma,t+1}(x^{t+1},y^{t+1})}{ITE^{\Gamma,t}(x^t,y^t)}}_{(IMEC)} \cdot \underbrace{\left[\frac{ITE^{\Gamma,t}(x^t,y^t)}{ITE^{\Gamma,t+1}(x^t,y^t)} \cdot \frac{ITE^{\Gamma,t}(x^{t+1},y^{t+1})}{ITE^{\Gamma,t+1}(x^{t+1},y^{t+1})}\right]^{1/2}}_{(IMTC)}, \tag{19}$$

where the ratio outside the brackets represents the relative input-oriented metatechnology efficiency change (*IMEC*) from period t to $t+1$. The part inside the brackets captures the shift of the metafrontiers between two periods. Due to the treatment of avoiding an arbitrary selection among years in (18), the geometric mean of the two ratios evaluated at (x^t, y^t) and (x^{t+1}, y^{t+1}) is specified as the input-oriented metatechnology change (*IMTC*).

Referring to (11), *IMEC* can be further represented in terms of *IMR* and *RITE*, which can be written:

$$IMEC^{\Gamma,t,t+1}(x^t,y^t,x^{t+1},y^{t+1}) = \underbrace{\frac{RITE^{g,t+1}(x^{t+1},y^{t+1})}{RITE^{g,t}(x^t,y^t)}}_{(ITEC)} \cdot \underbrace{\frac{IMR^{g\Gamma,t+1}(x^{t+1},y^{t+1})}{IMR^{g\Gamma,t}(x^t,y^t)}}_{(IMRC)}, \tag{20}$$

where the first part measures the efficiency changes with respect to the technology-specific frontier (i.e., input-oriented technology-specific efficiency change (*ITEC*)), whereas the second part depicts the input-oriented change of *IMR* (*IMRC*) between periods t to $t+1$. The latter describes whether the distance between technology-specific frontier and metafrontier in period $t+1$ is getting smaller or larger than that in period t.

In conclusion, the proposed *IMMI* is represented by means of the following decomposition:

$$IMMI = ITEC \cdot IMRC \cdot IMTC. \tag{21}$$

Specifically, $ITEC > 1$ (< 1) indicates that the unit under evaluation is approaching (moving away from) the corresponding technology-specific frontier from period t to $t+1$ in the input-orientation. $IMRC > 1$ (< 1) shows that the technology-specific frontier is approaching (moving away from) its metafrontier from period t to $t+1$ in the input-orientation. Finally, the last component implies a metatechnology progress (regress) from period t to $t+1$ if $IMTC > 1$ (< 1).

Note that the *IMMI* and its components can all be affected by the convexification strategy applied to the computation of the metatechnology, except the *ITEC* component that is evaluated with respect to the group-specific frontiers only. Furthermore, note that more elaborate decompositions of the metafrontier Malmquist index have been proposed in the literature: see, e.g., Chen and Yang (2011).

4.2. Metafrontier Hicks–Moorsteen productivity index

The seminal article by Bjurek (1996) introduces a Hicks–Moorsteen TFP index with a base period t as the ratio of a Malmquist type output quantity index (*MO*) in base period t over a Malmquist type input quantity index (*MI*) in the same base period t. Applied to a metatechnology Γ, this period t based metatechnology Hicks–Moorsteen productivity (*MHM*) index boils down to the following:

$$MHM^{\Gamma,t}(x^t,y^t,x^{t+1},y^{t+1}) = \frac{MO^{\Gamma,t}(x^t,y^t,y^{t+1})}{MI^{\Gamma,t}(x^t,x^{t+1},y^t)}, \tag{22}$$

with

$$MO^{\Gamma,t}(x^t,y^t,y^{t+1}) = \frac{OTE^{\Gamma,t}(x^t,y^t)}{OTE^{\Gamma,t}(x^t,y^{t+1})}$$

and

$$MI^{\Gamma,t}(x^t,x^{t+1},y^t) = \frac{ITE^{\Gamma,t}(x^t,y^t)}{ITE^{\Gamma,t}(x^{t+1},y^t)}.$$

A metatechnology Hicks–Moorsteen productivity index larger (smaller) than unity indicates a gain (loss) in productivity.

Similarly, the period $t+1$ based metatechnology Hicks–Moorsteen TFP index is defined as follows:

$$MHM^{\Gamma,t+1}(x^t,y^t,x^{t+1},y^{t+1}) = \frac{MO^{\Gamma,t+1}(x^{t+1},y^{t+1},y^t)}{MI^{\Gamma,t+1}(x^t,x^{t+1},y^{t+1})}, \tag{23}$$

[4] A similar comparison using difference-based indicators rather than ratio-based indices is available in the literature: see, e.g., Kerstens, Shen, and Van de Woestyne (2018).

with

$$MO^{\Gamma,t+1}(x^{t+1},y^{t+1},y^t)=\frac{OTE^{\Gamma,t+1}(x^{t+1},y^t)}{OTE^{\Gamma,t+1}(x^{t+1},y^{t+1})}$$

and

$$MI^{\Gamma,t+1}(x^t,x^{t+1},y^{t+1})=\frac{ITE^{\Gamma,t+1}(x^t,y^{t+1})}{ITE^{\Gamma,t+1}(x^{t+1},y^{t+1})}.$$

Again, a metatechnology Hicks–Moorsteen productivity index larger (smaller) than unity points to a productivity gain (loss).

To avoid an arbitrary choice of base year, it is customary to take a geometric mean of these two Hicks–Moorsteen TFP indices (22) and (23) (see Bjurek, 1996):

$$MHM^{\Gamma,t,t+1}(x^t,y^t,x^{t+1},y^{t+1}) = \left[MHM^{\Gamma,t}(x^t,y^t,x^{t+1},y^{t+1})\cdot MHM^{\Gamma,t+1}(x^t,y^t,x^{t+1},y^{t+1})\right]^{1/2}. \quad (24)$$

Note once more that a geometric mean metatechnology Hicks–Moorsteen productivity index larger (smaller) than unity indicates a productivity gain (loss).

Note that decompositions of the Hicks–Moorsteen TFP index are still rare in the literature. A recent proposal for a decomposition is the one by Diewert and Fox (2017), but it has rarely if ever been applied. The sole study applying the metatechnology Hicks–Moorsteen index (i.e., Verschelde et al., 2016) did not decompose. The comparison study of the Malmquist and Hicks–Moorsteen indices with standard technologies of Kerstens and Van de Woestyne (2014) did not decompose as well. Therefore, we do not develop a decomposition of the Hicks–Moorsteen index here.

Note that under specific assumptions on the group technologies these two productivity indices coincide: see Färe, Grosskopf, and Roos (1996; 1998), Bjurek, Førsund, and Hjalmarsson (1998), and O'Donnell (2012) for more details, and Kerstens and Van de Woestyne (2014, Section 2.4) for a brief review.

5. Nonparametric frontier technologies

If each TPPS is convex and exhibits VRS at time period t, then this TPPS is defined as:

$$T^{g,t}_{C,VRS}=\left\{(x^t,y^t)\in\mathbb{R}^M_+\times\mathbb{R}^N_+ : \sum_{i=1}^{n^g}\lambda_{\phi_g(i)}x^t_{\phi_g(i)}\le x^t,\ \sum_{i=1}^{n^g}\lambda_{\phi_g(i)}y^t_{\phi_g(i)}\ge y^t,\ \sum_{i=1}^{n^g}\lambda_{\phi_g(i)}=1,\ \lambda_{\phi_g(i)}\in\mathbb{R}_+\right\}, \quad (25)$$

with $s^{g,t}=\{(x^t_{\phi_g(i)},y^t_{\phi_g(i)}) : i=1,\dots,n^g\}$ the set of n^g initial observations at time period t determining technology g. The associated MTPPS is simply defined as the union of the above TPPS:

$$T^{\Gamma,t}_{C,VRS}=\cup_{g\in\Gamma}T^{g,t}_{C,VRS}. \quad (26)$$

The convexified version of (26) yields the following metatechnology:

$$H^{\Gamma,t}_{C,VRS}=\left\{(x^t,y^t)\in\mathbb{R}^M_+\times\mathbb{R}^N_+ : \sum_{g\in\Gamma}\sum_{i=1}^{n^g}\lambda_{\phi_g(i)}x^t_{\phi_g(i)}\le x^t,\ \sum_{g\in\Gamma}\sum_{i=1}^{n^g}\lambda_{\phi_g(i)}y^t_{\phi_g(i)}\ge y^t,\ \sum_{g\in\Gamma}\sum_{i=1}^{n^g}\lambda_{\phi_g(i)}=1,\ \lambda_{\phi_g(i)}\in\mathbb{R}_+\right\}. \quad (27)$$

O'Donnell et al. (2008, p. 238) employ this specification to determine an estimate of metatechnology-specific technical efficiency.[5]

[5] In fact, O'Donnell et al. (2008) compute an output-oriented metatechnology-specific technical efficiency under the assumption that there is no technical change.

However, note that in general $T^{\Gamma,t}_{C,VRS}\subseteq H^{\Gamma,t}_{C,VRS}$ where equality holds only for restrictive special cases (e.g., if only one group technology exists), as follows from Proposition 5.5 in Kerstens et al. (2019). The basic question we address in this contribution is whether the use of (27) leads to efficiency estimates that are close to the estimates obtained using the unbiased estimator (26).

If each TPPS is convex and exhibits CRS at time period t, then this TPPS is defined as:

$$T^{g,t}_{C,CRS}=\left\{(x^t,y^t)\in\mathbb{R}^M_+\times\mathbb{R}^N_+ : \sum_{i=1}^{n^g}\lambda_{\phi_g(i)}x^t_{\phi_g(i)}\le x^t,\ \sum_{i=1}^{n^g}\lambda_i y^t_{\phi_g(i)}\ge y^t,\ \lambda_{\phi_g(i)}\in\mathbb{R}_+\right\}, \quad (28)$$

with $s^{g,t}=\{(x^t_{\phi_g(i)},y^t_{\phi_g(i)}) : i=1,\dots,n^g\}$ the set of n^g initial observations at time period t determining technology g. The associated MTPPS is again defined as the union of the previous TPPS:

$$T^{\Gamma}_{C,CRS}=\cup_{g\in\Gamma}T^{g,t}_{C,CRS}. \quad (29)$$

The convexified version of (29) defines the following metatechnology:

$$H^{\Gamma,t}_{C,CRS}=\left\{(x^t,y^t)\in\mathbb{R}^M_+\times\mathbb{R}^N_+ : \sum_{g\in\Gamma}\sum_{i=1}^{n^g}\lambda_{\phi_g(i)}x^t_{\phi_g(i)}\le x^t,\ \sum_{g\in\Gamma}\sum_{i=1}^{n^g}\lambda_{\phi_g(i)}y^t_{\phi_g(i)}\ge y^t,\ \lambda_{\phi_g(i)}\in\mathbb{R}_+\right\}. \quad (30)$$

Proposition 5.5 of Kerstens et al. (2019) implies that $T^{\Gamma,t}_{C,CRS}\subseteq H^{\Gamma,t}_{C,CRS}$, whereby equality only holds for restrictive special cases. Again, the basic question is whether the use of (30) generates efficiency estimates that are close to the estimates obtained using the unbiased estimator (29).

The convex technology specifications with variable (i.e., (25) and (27)) and constant (i.e., (28) and (30)) returns to scale are commonly known as data envelopment analysis (moniker DEA) models. If the TPPS exhibit VRS (respectively CRS), then the technology specification (25) (respectively (28)) can be used to compute the measure of *RITE* (9) by solving a linear program for each evaluated observation (see Hackman, 2008 or Ray, 2004). The associated MTPPS (26) and (29) can be used to compute the measure of *ITE* (7) by solving for each evaluated observation several linear programs: one per TPPS. Recently, Afsharian and Podinovski (2018) show how to achieve the measure of *ITE* (7) relative to the MTPPS by solving a single LP problem. The convexification strategy embodied in the technologies (27) and (30) normally leads to a biased estimator for the *ITE* measure: it is an open empirical question how biased these computations exactly are when computing productivity indices with respect to a metafrontier.

Fig. 2 illustrates the issue at stake using variable returns to scale technologies (just as in Fig. 1). The TPPS $T^{1,t}$ and $T^{2,t}$ make up the associated MTPPS in period t. Similarly, the TPPS $T^{1,t+1}$ and $T^{2,t+1}$ constitute the components of the associated MTPPS in period $t+1$. From the discussion in Section 4 we recall that *IMMI* as well as its components can be affected by the convexification strategy applied to the MTTPS except for the *ITEC* component (since it is based on *RITE*). For instance, when projecting observation D_2 in period t (D^t_2) to the MTTPS in period t, then we can either project on the true MTTPS at point D'_2 or at the convexified MTTPS at point D''_2. When projecting this same observation to the MTTPS in period $t+1$, then we can either project on the true MTTPS at point D'''_2 or at the convexified MTTPS at point D''''_2. Obviously, these measurements involving *ITE* may affect both the *IMRC* and *IMTC* components.

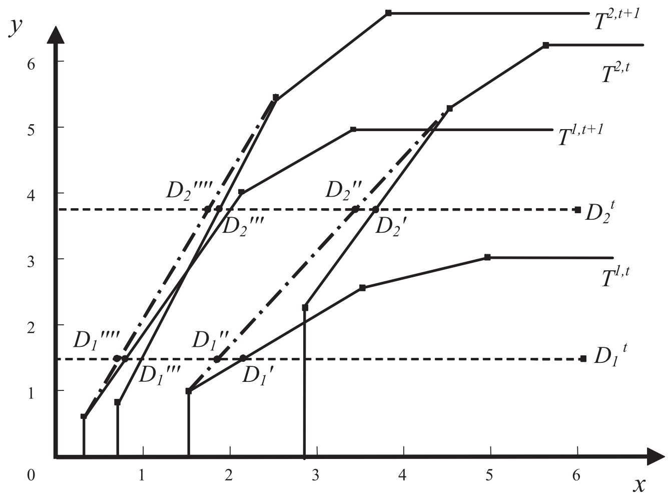

Fig. 2. Group technologies and metatechnologies over time.

6. Empirical illustration: hydroelectric power plants

6.1. Secondary data

This empirical section aims to illustrate the implications of a convexification strategy using secondary data that were previously used by Atkinson and Dorfman (2009) to evaluate the performance of an unbalanced panel of Chilean hydroelectric power plants. These data are publicly available on the data repository of the *Journal of Applied Econometrics*.[6] This sample comprises monthly data on $M = 3$ inputs and $N = 1$ output for 21 Chilean hydroelectric power plants over the period from April 1986 to December 1997. There are 7 dam plants and 14 run-of-river (ROR) plants in this sample. The three inputs are labor (in thousands of workers), capital (in real pesos), and water (in cubic meters). The single output is electricity generated (in gigawatt hours). More details regarding these data is found in Atkinson and Dorfman (2009) and Atkinson and Halabí (2005).

In Chile two main techniques (technologies) are employed to generate hydroelectric power. The first technology (index 1) builds a dam on a river to store water and releases this water from the dam to spin turbines generating electricity. The main advantage of dam systems is that electricity generation is uncoupled from river flows. The second hydroelectric power technology (index 2) involves merely diverting river flows through turbines. The advantage of these ROR or diversion systems is that these are relatively inexpensive and have relatively little impact on the environment. A key disadvantage of such systems is that these cannot be used to match electricity generation with consumer demand.[7] By construction, the technology set $\Gamma = \{1, 2\}$.

Our understanding of hydroelectric power generation leads us to believe that it *may* be possible for the manager of a given dam (ROR) plant to use a given input vector to produce a given level of output for some time within the planning period, and then use a different input vector to produce a different level of output for the rest of the time. This suggests that each TPPS *may* be convex. Consequently, we begin by computing these TPPSs t^1 and t^2 using a convex nonparametric frontier technology. Given the different types of capital involved in constructing different plants, it is also our understanding that the manager of a given plant cannot learn how to improve the performance by convex combinations of dam systems and ROR systems. This suggests that the MTPPS should not be convexified. It is now an open question to check how a convexification strategy of the MTPPS approximates the true nonconvex MTPPS.

6.2. Empirical results

Table 1 contains basic descriptive statistics of the *IMMI* estimates and its components with balanced and unbalanced panel data of hydroelectric power plants under CRS and VRS technologies, respectively. This table is structured as follows: first we discuss the columns, then we explain the rows. The first four columns list the results under a CRS technology and the last four columns assume a VRS technology. Within each of these technologies, the first two columns report the results with unbalanced panel data while the last two columns develop the balanced panel data results. A further distinction is related to whether a convexification strategy is applied or not: a C indicates a convexification strategy is applied to calculate *ITE*, while NC reveals that this is not the case. Horizontally, the first block of rows contains the results of the *IMMI* estimates. The following three horizontal blocks of rows reports the decomposition results of *IMMI*. Within each of these four horizontal blocks, we report results on geometric means, standard deviations, minimum and maximum values of the corresponding estimates. The use of geometric means guarantees the multiplicative decomposition of the *IMMI*.

Furthermore, a nonparametric Li-test is applied to test the null hypothesis that the distributions of the two C vs. NC *IMMI* as well as its two components *IMRC* and *IMTC* are equal. Since the *ITEC* component does not differ under C vs. NC, no test statistic is computed. This Li-test is first proposed by Li (1996) and has been refined by Fan and Ullah (1999) and others: one of the most recent developments is by Li, Maasoumi, and Racine (2009). This nonparametric test analyzes the differences between entire distributions by comparing the differences between two kernel-based estimates of density functions f and g of a random variable x. The null hypothesis affirms that both density functions are almost everywhere equal ($H_0 : f(x) = g(x)$ for all x). The alternative hypothesis simply negates this equality of both density functions (H_1: $f(x) \neq g(x)$ for some x). Simar and Zelenyuk (2006) refine this test statistic further for nonparametric frontier estimators to circumvent the problem of the potential spurious mass at the boundary. While this is a general problem for efficiency measures, the productivity indices as ratios of efficiency measures do not suffer from this problem. The test statistics marked with "***" means the null hypothesis is rejected at the 0.1% significance level. A large (resp. small) p value indicates that the null hypothesis should not be rejected (resp. be rejected).[8]

Finally, the number as well as the percentage of contradictory results implied by comparing estimates under C and NC strategies are reported in the last row of each block (denoted "#Contrad. Res."). Contradictory results arise when one estimate points to a productivity decline while the other shows a productivity growth, or the other way around.

Note that the computation of these descriptive statistics and tests is based on the productivity indicators available. Hence, the number of valid results may differ from case to case. Taking the case of balanced panel data as an example, there are 1085 valid *IMMI* results under CRS. Due to the occurrence of computational infeasibilities, the valid number under VRS is only 933.

[6] Web site: http://qed.econ.queensu.ca/jae/.

[7] A third hydroelectric power technology, called pumped storage, allows to match electricity generation with variations in consumer demand. But, there are no pumped storage plants in this sample.

[8] The Matlab code for the Li-test adopted here is developed by P.J. Kerstens based on Li et al. (2009). The Simar and Zelenyuk (2006) refinement is also an option. This code is found at: https://github.com/kepiej/DEAUtils.

Table 1
Descriptive statistics and Li-test for the estimates of *IMMI* and its decompositions.

		CRS				VRS			
		Unbalanced		Balanced		Unbalanced		Balanced	
		C	NC	C	NC	C	NC	C	NC
IMMI	Mean	1.0434	1.0443	1.0181	1.0191	1.0397	1.0461	1.0246	1.0361
	Std. dev.	0.3486	0.3492	0.1911	0.1982	0.3084	0.3321	0.2348	0.3106
	Min	0.1687	0.1660	0.3970	0.3970	0.1853	0.1662	0.1859	0.2163
	Max	6.1295	6.0020	2.4833	2.4833	5.0747	5.3552	2.6642	2.6642
	Li-test		−2.4605		0.5387		4.3619***		7.3193***
	p-value		(0.999)		(0.2675)		(0.001)		(0.0005)
	#Inf.		0/2412		0/1085		185/2412		152/1085
	Res.		(0)		(0)		(7.67%)		(14.01%)
	#Contrad.		35/2412		29/1085		200/2227		73/933
	Res.		(1.45%)		(2.67%)		(8.98%)		(7.82%)
ITEC	Mean	1.0245		1.0028		1.0121		1.0013	
	Std. dev.	0.2616		0.0774		0.1924		0.0531	
	Min	0.1705		0.5906		0.2134		0.5672	
	Max	4.3616		1.7857		5.0214		1.6507	
IMRC	Mean	1.0105	1.0099	1.0108	1.0116	1.0104	1.0061	1.0090	1.0040
	Std. dev.	0.1694	0.1692	0.1516	0.1581	0.1561	0.1374	0.1376	0.0861
	Min	0.3107	0.2944	0.4916	0.4916	0.4072	0.2379	0.5020	0.5118
	Max	5.3189	5.7023	1.9405	1.9405	2.4700	4.2052	2.1427	2.0855
	Li-test		483.4166***		259.9905***		158.8927***		489.1541***
	p-value		(0.0000)		(0.0000)		(0.0000)		(0.0000)
	#Contrad.		731/2412		230/1085		873/2227		268/933
	Res.		(30.31%)		(21.20%)		(39.20%)		(28.72%)
IMTC	Mean	1.0130	1.0143	1.0119	1.0110	1.0211	1.0331	1.0141	1.0269
	Std. dev.	0.1431	0.1472	0.1480	0.1440	0.1859	0.2530	0.1751	0.2742
	Min	0.5246	0.5215	0.5246	0.5215	0.2157	0.2220	0.1962	0.2163
	Max	2.0983	2.1508	1.8789	1.8789	2.7293	3.1923	2.6642	2.6642
	Li-test		−2.0476		1.0960		6.3767***		6.5930***
	p-value		(0.9930)		(0.1245)		(0.0000)		(0.0000)
	#Contrad.		127/2412		57/1085		349/2227		153/933
	Res.		(5.27%)		(5.25%)		(15.67%)		(16.40%)

Several observations can be made with regard to the results in Table 1. First, the basic descriptive statistics for *IMMI*, *IMRC* and *IMTC* estimates all show certain differences when comparing between C and NC strategies. In addition, the *IMMI* estimates under a C strategy are on average lower than the ones under a NC strategy in our data. Theoretically, estimates of these indices derived under C strategy can be either higher or lower than the ones derived under NC strategy. This is consistent with the observations for the *IMRC* and *IMTC* estimates. As for the *ITEC* estimates, which capture the efficiency changes with respect to technology-specific frontiers, the convexification strategy makes no difference.

Second, the Li-test for the *IMRC* estimates reveals that the C strategy leads to a significant difference in the distribution of the metafrontier compared with the NC strategy. This holds true for both CRS and VRS technologies. For the VRS technology, a statistically significant difference can be found for all *IMMI* and *IMTC* estimates. Thus, for our data the convexification strategy has a stronger influence on calculating these estimates under VRS than under CRS.

Third, all estimates under various cases offer some contradictory signs between applying C and NC strategies. On average, more contradictory signs are detected in the estimates of *IMRC* and *IMTC* than that of *IMMI*. The percentage of contradictory signs reaches 39.2% for *IMRC* estimates and 16.40% for *IMTC*. The contradictory signs appear more frequently under the VRS technology than under CRS. This coincides with the above finding that the bias of applying the convexification strategy is more evident under VRS than under CRS for our data.

The descriptive statistics results, results for the nonparametric Li-test, and contradictory signs for the metafrontier Hicks–Moorsteen (*MHM*) TFP estimates are displayed in Table 2. This table is similar in structure to Table 1, except that only the *MHM* estimate is reported and no decomposition.

The following observations can be made regarding this Table 2. First, minor differences are observed between the C and NC strategies from the basic descriptive statistics of the *MHM* estimates. Second, a statistically significant difference in distributions is detected by the Li-test for the case with balanced data and under the CRS assumption only. Third, contradictory results exist while comparing between C and NC strategies for all four cases. Furthermore, more opposite signs show up under the VRS assumption than under the CRS assumption. Fourth, no infeasibilities are recorded (see Briec & Kerstens, 2011 who prove that the Hicks–Moorsteen index does not yield any infeasibilities under standard assumptions on technology).

In general, we can conclude that applying a convexification strategy for both productivity indices shows quite a difference from the original non-convex metafrontier productivity indices. The contradictory results also underscore the drawback of applying a convexification strategy. More specifically, there is a non-negligible possibility that the suggestions based on the estimates obtained by applying a convexification strategy lead to opposite conclusions and policy recommendations.

In Table 3 the degree of similarity between the metafrontier Malmquist and Hicks–Moorsteen productivity indices – both obtained under the correct NC strategy – is investigated. Although a detailed analysis of this similarity exists in the literature at the normal frontier level (see, e.g., Kerstens & Van de Woestyne, 2014), their similarity at the metafrontier level has – to the best of our knowledge – not been compared before. First, note that the empirical study here contains a single output and multiple inputs. This makes both metafrontier productivity indices coincide under

Table 2
Descriptive statistics and Li-test for the estimate of *MHM*.

	CRS				VRS			
	Unbalanced		Balanced		Unbalanced		Balanced	
	C	NC	C	NC	C	NC	C	NC
Mean	1.0434	1.0443	1.0181	1.0191	1.0496	1.0489	1.0254	1.0262
Std. dev.	0.3486	0.3492	0.1911	0.1982	0.4033	0.3924	0.2484	0.2586
Min	0.1687	0.1660	0.3970	0.3970	0.1040	0.1040	0.1774	0.1774
Max	6.1295	6.0020	2.4833	2.4833	8.3617	8.0871	4.6910	4.6910
Li-test		−2.4605		9.3788***		−2.6610		−1.9210
p-value		(1.0000)		(0.0000)		(1.0000)		(0.9995)
# Inf.		0/2412		0/1085		0/2412		0/1085
Res.		(0)		(0)		(0)		(0)
# Contrad.		31/2412		29/1085		52/2412		39/1085
Res.		(1.29%)		(2.67%)		(2.16%)		(3.59%)

Table 3
Descriptive statistics and Li-test for the estimates of *IMMI* and *MHM*.

	VRS			
	Unbalanced		Balanced	
	IMMI	MHM	IMMI	MHM
Mean	1.0461	1.0489	1.0361	1.0262
Std. dev.	0.3321	0.3924	0.3106	0.2586
Min	0.1662	0.1040	0.2163	0.1774
Max	5.3552	8.0871	2.6642	4.6910
Li-test		3.8946**		13.7819***
p-value		(0.0020)		(0.0000)
# Inf.	185/2412	0/2412	152/1085	0/1085
Res.	(7.67%)	(0)	(14.01%)	(0)
# Contrad.		281/2227		145/933
Res.		(12.62%)		(15.54%)

the CRS assumption (see Bjurek et al., 1998). However, both indices potentially remain to show differences under VRS. Second, we can observe that all descriptive statistics somewhat vary between *IMMI* and *MHM*. There are no infeasibilities detected in calculating *MHM*, while infeasibilities arise in calculating *IMMI* under VRS. Accordingly, the results where the distributions and contradictory signs are examined are based on available indices. Third, the Li-test shows that the distributions of *IMMI* and *MHM* are significantly different for the balanced data under VRS at the significance level of 0.1%. For the case with unbalanced data under VRS, their distributions are significantly different at the significance level of 1%, which is marked with "**" in Table 3. Finally, the percentage having contradictory results between *IMMI* and *MHM* reaches 15.54% and 12.62% under balanced and unbalanced panel data, respectively. Therefore, the above observations all imply that the *IMMI* and *MHM* indices under VRS are empirically distinct.

The main purpose of this empirical application is twofold. First, we investigate the impact of a convexification strategy on the estimation of metafrontier-based productivity indices. The second goal is to compare the differences between Malmquist and Hicks–Moorsteen indices at the metafrontier level. For these purposes, the findings of descriptive statistics, Li-tests and contradictory signs are presented as analytical findings to show statistical evidence on the bias of a convexification strategy and the differences between both indices.

7. Conclusions

In their seminal article, O'Donnell et al. (2008) define a nonconvex metatechnology as the union of two or more underlying group-specific technologies. They suggest estimating the metafrontier (i.e., the boundary of the metatechnology) under the assumption that the metatechnology is convex. If this assumption yields a poor approximation of the true nonconvex metatechnology, then their convexification strategy yields a biased estimator. Recently, Kerstens et al. (2019) develop some new results on the union operation on technologies under various assumptions of returns to scale and convexity and are the first to statistically test that a convexification strategy leads to erroneous results. Note that this controversy exists when the group-specific frontier is convex. But, as shown in Verschelde et al. (2016) and Walheer (2018), the group-specific frontiers could also be estimated relative to a nonconvex setting. Then, the corresponding metafrontier is again naturally nonconvex. Prior to choosing specific estimates, practitioners may explore the proper group-specific technology (e.g., returns to scale or convexity) in the way suggested by Kneip, Simar, and Wilson (2016).

This is the first empirical contribution that explores the impact of a convexification strategy on a Malmquist productivity index evaluated relative to a metafrontier methodology. Theoretically, we find that the input-oriented Malmquist productivity index itself as well as its components input-oriented metatechnology change (*IMTC*) and input-oriented metatechnology ratio change (*IMRC*) are potentially affected by the choice of a convexification strategy. However, this is not the case for the input-oriented technical efficiency change (*ITEC*) component. Empirically, our results for a secondary date set reveal that a convexification strategy leads to statistically significant differences for the input-oriented Malmquist productivity index and its components. Furthermore, at the level of the individual observations it can lead to opposite signs for a substantial fraction of the sample.

Equally so, this is the first contribution that investigates the impact of a convexification strategy on a metafrontier Hicks–Moorsteen TFP index. Though at the level of the individual observations, we can observe some opposite signs for some fraction of the sample, our empirical results show that a convexification strategy only leads to statistically significant differences for the balanced CRS case.

Finally, to the best of our knowledge, this is the first study comparing the Malmquist technology index with the Hicks–Moorsteen TFP index in a metafrontier setting. Just as in the case of a standard technology, the CRS case in our setting being identical, we do find contradictory results and statistically significant differences for the VRS case. This confirms earlier comparative results on standard technologies (e.g., Kerstens & Van de Woestyne, 2014).

In conclusion, we can state that this contribution has shown that a convexification strategy threatens to undermine the metafrontier methodology by yielding biased results when the group technologies are assumed to be convex. We may safely assume that these conclusions also transpose to alternative productivity indices such as, e.g., the Luenberger indicator and the dual

Malmquist index, and to the primal TFP indices (such as the Färe–Primont and Lowe indices, among others). However, it would be good if future research sheds more light on this conjecture. There also remain other open challenges for future investigation. One interesting extension is related to inferential issues. While inference for Malmquist indices has been extensively studied (e.g., Simar & Wilson, 1999 or more recently Kneip, Simar, & Wilson, 2018), there is no closed analytical form for inference about the productivity changes measured by metafrontier productivity indices.

Acknowledgment

We thank three referees for most useful comments. The usual disclaimer applies. Q. Jin acknowledges the financial support of the China Scholarship Council for funding her stay in Lille.

References

Afsharian, M., & Ahn, H. (2015). The overall Malmquist index: A new approach for measuring productivity changes over time. *Annals of Operations Research, 226*(1), 1–27.

Afsharian, M., Ahn, H., & Harms, S. (2018). A non-convex meta-frontier Malmquist index for measuring productivity over time. *IMA Journal of Management Mathematics, 29*(4), 377–392.

Afsharian, M., & Podinovski, V. (2018). A linear programming approach to efficiency evaluation in nonconvex metatechnologies. *European Journal of Operational Research, 268*(1), 268–280.

Atkinson, S., & Dorfman, J. (2009). Feasible estimation of firm-specific allocative inefficiency through Bayesian numerical methods. *Journal of Applied Econometrics, 24*(4), 675–697.

Atkinson, S., & Halabí, C. (2005). Economic efficiency and productivity growth in the post-privatization Chilean hydroelectric industry. *Journal of Productivity Analysis, 23*(2), 245–273.

Balk, B. (2016). Various approaches to the aggregation of economic productivity indices. *Pacific Economic Review, 21*(4), 445–463.

Battese, G., & Rao, D. (2002). Technology gap, efficiency and a stochastic metafrontier function. *International Journal of Business and Economics, 1*(2), 87–93.

Battese, G., Rao, D., & O'Donnell, C. (2004). A metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies. *Journal of Productivity Analysis, 21*(1), 91–103.

Bjurek, H. (1996). The Malmquist total factor productivity index. *Scandinavian Journal of Economics, 98*(2), 303–313.

Bjurek, H., Førsund, F. R., & Hjalmarsson, L. (1998). Malmquist productivity indexes: An empirical comparison. In *Index numbers: Essays in honour of sten Malmquist* (pp. 217–239). Springer.

Briec, W., & Kerstens, K. (2009). Infeasibilities and directional distance functions with application to the determinateness of the Luenberger productivity indicator. *Journal of Optimization Theory and Applications, 141*(1), 55–73.

Briec, W., & Kerstens, K. (2011). The Hicks–Moorsteen productivity index satisfies the determinateness axiom. *Manchester School, 79*(4), 765–775.

Caselli, F., & Coleman, W. (2006). The world technology frontier. *American Economic Review, 93*(3), 499–522.

Casu, B., Ferrari, A., & Zhao, T. (2013). Regulatory reform and productivity change in Indian banking. *Review of Economics and Statistics, 95*(3), 1066–1077.

Caves, D., Christensen, L., & Diewert, W. (1982). The economic theory of index numbers and the measurement of inputs, outputs and productivity. *Econometrica, 50*(6), 1393–1414.

Chen, K.-H., & Yang, H.-Y. (2011). A cross-country comparison of productivity growth using the generalised metafrontier Malmquist productivity index: With application to banking industries in Taiwan and China. *Journal of Productivity Analysis, 35*(3), 197–212.

Dakpo, K., Desjeux, Y., Jeanneaux, P., & Latruffe, L. (2019). Productivity, technical efficiency and technological change in French agriculture during 2002–2015: A Färe-Primont index decomposition using group frontiers and meta-frontier. *Applied Economics, 51*(11), 1166–1182.

Diewert, W., & Fox, K. (2017). Decomposing productivity indexes into explanatory factors. *European Journal of Operational Research, 256*(1), 275–291.

Fan, Y., & Ullah, A. (1999). On goodness-of-fit tests for weakly dependent processes using kernel method. *Journal of Nonparametric Statistics, 11*(1), 337–360.

Färe, R., Grosskopf, S., Lindgren, B., & Roos, P. (1992). Productivity changes in Swedish pharamacies 1980–1989: A non-parametric Malmquist approach. *Journal of Productivity Analysis, 3*(1–2), 85–101.

Färe, R., Grosskopf, S., Lindgren, B., & Roos, P. (1995). Productivity developments in Swedish hospitals: A Malmquist output index approach. In A. Charnes, W. Cooper, A. Lewin, & L. Seiford (Eds.), *Data envelopment analysis: Theory, methodology and applications* (pp. 253–272). Boston: Kluwer.

Färe, R., Grosskopf, S., Norris, M., & Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *American Economic Review, 84*(1), 66–83.

Färe, R., Grosskopf, S., & Roos, P. (1996). On two definitions of productivity. *Economics Letters, 53*(3), 269–274.

Färe, R., Grosskopf, S., & Roos, P. (1998). Malmquist productivity indexes: A survey of theory and practice. In *Index numbers: Essays in honour of sten Malmquist* (pp. 127–190). Springer.

Färe, R., & Primont, D. (1995). *Multi-output production and duality: Theory and applications*. Springer Science & Business Media.

Grosskopf, S. (2003). Some remarks on productivity and its decompositions. *Journal of Productivity Analysis, 20*(3), 459–474.

Hackman, S. (2008). *Production economics: Integrating the microeconomic and engineering perspectives*. Berlin: Springer.

Hayami, Y., & Ruttan, V. (1970). Agricultural productivity differences among countries. *American Economic Review, 60*(5), 895–911.

Huang, C.-W., Ting, C.-T., Lin, C.-H., & Lin, C.-T. (2013). Measuring non-convex metafrontier efficiency in international tourist hotels. *Journal of the Operational Research Society, 64*(2), 250–259..

Huang, M.-Y., & Fu, T.-T. (2013). An examination of the cost efficiency of banks in Taiwan and China using the metafrontier cost function. *Journal of Productivity Analysis, 40*(3), 387–406.

Huang, M.-Y., Juo, J.-C., & Fu, T.-T. (2015). Metafrontier cost Malmquist productivity index: An application to Taiwanese and Chinese commercial banks. *Journal of Productivity Analysis, 44*(3), 321–335.

Kerstens, K., O'Donnell, C., & Van de Woestyne, I. (2019). Metatechnology frontier and convexity: A restatement. *European Journal of Operational Research, 275*(2), 780–792.

Kerstens, K., Shen, Z., & Van de Woestyne, I. (2018). Comparing Luenberger and Luenberger–Hicks–Moorsteen productivity indicators: How well is total factor productivity approximated? *International Journal of Production Economics, 195*, 311–318.

Kerstens, K., & Van de Woestyne, I. (2014). Comparing Malmquist and Hicks—Moorsteen productivity indices: Exploring the impact of unbalanced vs. balanced panel data. *European Journal of Operational Research, 233*(3), 749–758.

Kneip, A., Simar, L., & Wilson, P. (2016). Testing hypotheses in nonparametric models of production. *Journal of Business & Economic Statistics, 34*(3), 435–456.

Kneip, A., Simar, L., & Wilson, P. W. (2018). Inference in dynamic, nonparametric models of production: Central limit theorems for Malmquist indices. *Technical Report Discussion paper 2018/10*. ISBA, UCL.

Kounetas, K., Mourtos, I., & Tsekouras, K. (2009). Efficiency decompositions for heterogeneous technologies. *European Journal of Operational Research, 199*(1), 209–218.

Latruffe, L., Fogarasi, J., & Desjeux, Y. (2012). Efficiency, productivity and technology comparison for farms in Central and Western Europe: The case of field crop and dairy farming in Hungary and France. *Economic Systems, 36*(2), 264–278.

Lau, L., & Yotopoulos, P. (1989). The meta-production function approach to technological change in world agriculture. *Journal of Development Economics, 31*(2), 241–269.

Lee, S.-G., & Midani, A. (2015). Comparison of efficiency levels using meta-frontier analysis of global fisheries for the period 1960–2010. *Fisheries Science, 81*(2), 247–254.

Li, Q. (1996). Nonparametric testing of closeness between two unknown distribution functions. *Econometric Reviews, 15*(1), 261–274.

Li, Q., Maasoumi, E., & Racine, J. (2009). A nonparametric test for equality of distributions with mixed categorical and continuous data. *Journal of Econometrics, 148*(2), 186–200.

Lovell, C. (2003). The decomposition of Malmquist productivity indexes. *Journal of Productivity Analysis, 20*(3), 437–458.

Mayer, A., & Zelenyuk, V. (2014). Aggregation of Malmquist productivity indexes allowing for reallocation of resources. *European Journal of Operational Research, 238*(3), 774–785.

O'Donnell, C. (2012). An aggregate quantity framework for measuring and decomposing productivity and profitability change. *Journal of Productivity Analyis, 38*(3), 255–272.

O'Donnell, C. (2016). Using information about technologies, markets and firm behaviour to decompose a proper productivity index. *Journal of Econometrics, 190*(2), 328–340.

O'Donnell, C. (2018). *Productivity and efficiency analysis: An economic approach to measuring and explaining managerial performance*. Singapore: Springer.

O'Donnell, C., Fallah-Fini, S., & Triantis, K. (2017). Measuring and analysing productivity change in a metafrontier framework. *Journal of Productivity Analysis, 47*(2), 117–128.

O'Donnell, C., Rao, D., & Battese, G. (2008). Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. *Empirical Economics, 34*(1), 231–255.

Orea, L., & Kumbhakar, S. (2004). Efficiency measurement using a latent class stochastic frontier model. *Empirical Economics, 29*(1), 169–183.

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

Russell, R. (2018). Theoretical productivity indices. In E. Grifell-Tatjé, C. Lovell, & R. Sickles (Eds.), *The Oxford handbook on productivity analysis* (pp. 153–182). New York: Oxford University Press.

Sala-Garrido, R., Molinos-Senante, M., & Hernández-Sancho, F. (2011). Comparing the efficiency of wastewater treatment technologies through a DEA metafrontier model. *Chemical Engineering Journal, 173*(3), 766–772.

Simar, L., & Wilson, P. W. (1999). Estimating and bootstrapping Malmquist indices. *European Journal of Operational Research, 115*(3), 459–471.

Simar, L., & Zelenyuk, V. (2006). On testing equality of distributions of technical efficiency scores. *Econometric Reviews, 25*(4), 497–522.

Thieme, C., Prior, D., & Tortosa-Ausina, E. (2013). A multilevel decomposition of school performance using robust nonparametric frontier techniques. *Economics of Education Review, 32*, 104–121.

Tiedemann, T., Francksen, T., & Latacz-Lohmann, U. (2011). Assessing the performance of German Bundesliga football players: A non-parametric metafrontier approach. *Central European Journal of Operations Research, 19*(4), 571–587.

Trueblood, M. (1989). Agricultural production function estimates from aggregate intercountry observations: A selected survey. *Canadian Journal of Agricultural Economics, 37*(4), 1045–1060.

Tsekouras, K., Chatzistamoulou, N., & Kounetas, K. (2017). Productive performance, technology heterogeneity and hierarchies: Who to compare with whom. *International Journal of Production Economics, 193*, 465–478.

Verschelde, M., Dumont, M., Rayp, G., & Merlevede, B. (2016). Semiparametric stochastic metafrontier efficiency of European manufacturing firms. *Journal of Productivity Analysis, 45*(1), 53–69.

Walheer, B. (2018). Aggregation of metafrontier technology gap ratios: The case of European sectors in 1995–2015. *European Journal of Operational Research, 269*(3), 1013–1026.

Zhang, N., & Wang, B. (2015). A deterministic parametric metafrontier Luenberger indicator for measuring environmentally-sensitive productivity growth: A Korean fossil-fuel power case. *Energy Economics, 51*, 88–98.

Zofio, J. L. (2007). Malmquist productivity index decompositions: A unifying framework. *Applied Economics, 39*(18), 2371–2387.