European Journal of Operational Research 233 (2014) 749–758

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Interfaces with Other Disciplines

Comparing Malmquist and Hicks–Moorsteen productivity indices: Exploring the impact of unbalanced vs. balanced panel data ☆

Kristiaan Kerstens [a,*], Ignace Van de Woestyne [b]

[a] CNRS-LEM (UMR 8179), IESEG School of Management, 3 rue de la Digue, F-59000 Lille, France
[b] Hogeschool-Universiteit Brussel, Warmoesberg 26, B-1000 Brussel, Belgium

ARTICLE INFO

Article history:
Received 31 December 2012
Accepted 9 September 2013
Available online 20 September 2013

Keywords:
Malmquist productivity index
Hicks–Moorsteen productivity index
Balanced panel
Unbalanced panel

ABSTRACT

We explore the effect of balancing unbalanced panel data when estimating primal productivity indices using non-parametric frontier estimators. First, we list a series of pseudo-solutions aimed at making an unbalanced panel balanced. Then, we discuss some intermediate solutions (e.g., balancing 2-years by 2 years). Furthermore, we link this problem with a variety of literatures on infeasibilities, statistical inference of non-parametric frontier estimators, and the index theory literature focusing on the dynamics of entry and exit in industries. We then empirically illustrate these issues comparing both Malmquist and Hicks–Moorsteen productivity indices on two data sets. In particular, we test for the differences in distribution when comparing balanced and unbalanced results for a given index and when comparing Malmquist and Hicks–Moorsteen productivity indices for a given type of data set. The latter tests are crucial in answering the question to which extent the Malmquist index can approximate the Hicks–Moorsteen index that has a Total Factor Productivity (TFP) interpretation. Finally, we draw up a list of remaining issues that could benefit from further exploration.

1. Introduction

Traditionally, Total Factor Productivity (TFP) growth is estimated by the traditional Solow residual and yields an index number representing technology shifts from output growth that remain unexplained by input growth (see Hulten (2001) or Van Beveren (2010)). In the last decades, economists have become conscious that ignoring inefficiency may well bias TFP measures. Nishimizu and Page (1982) is probably the seminal article suggesting to decompose TFP into a technical change component as well as a technical efficiency change component. Caves, Christensen, and Diewert (1982) have analyzed discrete time Malmquist input, output and productivity indices using distance functions as general technology representations. Since these Malmquist indices require a precise knowledge on the technology, these authors relate Malmquist and Törnqvist productivity indices, the latter depending on both price and quantity information (without need of exact knowledge on the technology).

Integrating the two-part Nishimizu and Page (1982) decomposition, Färe, Grosskopf, Norris, and Zhang (1994) propose to estimate the output distance functions in the Malmquist output productivity index by exploiting their inverse relation with the radial output efficiency measures evaluated relative to multiple input and output non-parametric technologies. Meanwhile, parametric estimates of the underlying distance functions of this Malmquist productivity index approach have also been reported (see, e.g., Atkinson, Cornwell, & Honerkamp (2003) or Tsekouras, Pantzios, & Karagiannis (2004)). Bjurek (1996) offers an alternative Hicks–Moorsteen TFP index, defined as a ratio of a Malmquist output over a Malmquist input index (see also O'Donnell (2010, 2012a)). Finally, it is good to indicate that these primal productivity indices have become relatively popular in empirical work in comparison with more traditional productivity measures (e.g., Fisher or Törnqvist indices).

This paper concentrates mainly on a seemingly rather widespread misconception that both these primal productivity indices require balanced panel data and cannot cope with unbalancedness. Just to cite one example, Hollingsworth and Wildman (2003) state that "DEA based Malmquist techniques are unable to cope with unbalanced panel estimation procedures" (page 497). One reason for such beliefs could be that some of the popular software options around to compute these productivity indices cannot handle unbalanced panels. For instance, the still popular DEAP software of Coelli (1996) explicitly requires a balanced panel (see p. 31 of the manual).[1] Such software restrictions may induce people to believe balanced panels are a prerequisite for this Malmquist

☆ We are grateful for generous comments offered by three referees. The usual disclaimer applies.

* Corresponding author.
E-mail address: k.kerstens@ieseg.fr (K. Kerstens).

[1] Another example of the same explicit requirement is the R-package "Nonparaeff" (version 0.5-6: page 14).

0377-2217/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2013.09.009

productivity index approach. This is to some extent surprising given that some of the seminal articles on the Malmquist productivity index have clearly pointed out that the use of an unbalanced panel is possible, "although the index will be undefined for missing observations" (see Färe et al. (1994, p. 73, fn 14)). While the notion of a potential unbalancedness bias due to unplanned missing data is quite standard in the statistical literature (see, e.g., Baltagi & Song (2006) or Frees (2004)), to the best of our knowledge nobody has so far analyzed the extent of the differences between computing primal productivity indices using balanced and unbalanced panel data.

A secondary goal of the paper is to formally test for the degree of similarity between the Malmquist and the Hicks–Moorsteen productivity indices. Following Bjurek, Førsund, and Hjalmarsson (1998), the recent literature has clearly established that the Malmquist productivity index has no TFP interpretation in general, while the Hicks–Moorsteen index does have a TFP interpretation (O'Donnell (2010, 2012a)). Only under rather stringent conditions on technology (see infra), both productivity indices coincide. To our knowledge this paper is the first to formally test whether both indices coincide empirically under a variety of technology specifications and in the presence of balanced or unbalanced panel data. Only when these indices turn out to be empirically indistinguishable, the Malmquist productivity index can maintain a TFP interpretation by approximation.

In this contribution, we intend to systematically start exploring the consequences of computing these primal productivity indices using a balanced panel when initially an unbalanced panel data set is available. We also formally test the degree of similarity between both productivity indices for a given structure of panel data. In particular, this paper is structured as follows. Section 2 provides some basic definitions of the technology, and of the Malmquist productivity index as well as the Hicks–Moorsteen TFP index. Section 3 offers a structured overview of different "solutions" advanced in the literature to cope with unbalanced panel data when computing these primal productivity indices. We argue against most of these pseudo-solutions. In Section 4, the effect of the balancedness or unbalancedness of the sample is illustrated using existing data sets. The final Section 5 concludes and outlines future research issues.

2. Definitions of technology and primal productivity indices

We first introduce the assumptions on technology and the definitions of the required distance functions. The latter provide the components for computing the primal productivity indices.

2.1. Technology and distance functions

A production technology describes how a vector of inputs $x = (x_1, \ldots, x_n) \in \mathbb{R}^n_+$ is transformed into a vector of outputs $y = (y_1, \ldots, y_p) \in \mathbb{R}^p_+$. For each time period t, the production possibility set (or technology for short) T^t summarizes the set of all feasible vectors of input and output. It is defined as follows:

$$T^t = \{(x^t, y^t) \in \mathbb{R}^{n+p}_+ : x^t \text{ can produce } y^t \text{ in period } t\}. \tag{1}$$

Throughout this contribution, technology is assumed to satisfy the following conventional assumptions:

(T.1) $(0,0) \in T^t$, $(0,y^t) \in T^t \Rightarrow y^t = 0$.
(T.2) The set $A(x^t) = \{(u^t, y^t) \in T^t: u^t \leqslant x^t\}$ of dominating observations is bounded $\forall x^t \in \mathbb{R}^n_+$.
(T.3) T^t is closed.
(T.4) $\forall (x^t, y^t) \in T^t$: $(x^t, -y^t) \leqslant (u^t, -v^t)$ and $(u^t, v^t) \geqslant 0$ implies that $(u^t, v^t) \in T^t$.

The first axiom creates the possibility of inaction and also states that there is no free lunch. The second axiom of boundedness (i.e., infinite outputs cannot be obtained from a finite input vector) is just a mathematical regularity condition, as is closedness of technology assumed in the third axiom. The fourth axiom of strong disposal of inputs and outputs implies that more inputs can always be used for given outputs and that fewer outputs can always be produced with given inputs.

Sometimes, the following two additional axioms are assumed in various combinations with the preceding ones as well:

(T.5) T^t is a convex set.
(T.6) $\delta T^t \subseteq T^t$, $\forall \delta > 0$.

Convexity of technology in the fifth axiom allows for linear combinations of activities to remain feasible. The sixth axiom imposes constant returns to scale rather than a more flexible variable returns to scale hypothesis that is traditionally maintained.

Efficiency is estimated relative to technologies using distance or gauge functions. Distance functions are related to the efficiency measures defined by Farrell (1957). In the input-orientation, this Farrell efficiency measure $E^i_t(x^t, y^t)$ indicates the minimum contraction of an input vector by a scalar λ while still remaining on the boundary of the technology:

$$E^i_t(x^t, y^t) = \inf_{\lambda}\{\lambda : (\lambda x^t, y^t) \in T^t, \lambda \geqslant 0\}. \tag{2}$$

In the output-orientation, the Farrell efficiency measure $E^o_t(x^t, y^t)$ searches for the maximum expansion of the output vector by a scalar θ to the boundary of the technology:

$$E^o_t(x^t, y^t) = \sup_{\theta}\{\theta : (x^t, \theta y^t) \in T^t, \theta \geqslant 1\}. \tag{3}$$

Following Färe et al. (1994, p. 69 (esp. fn 4)), the Farrell efficiency measure $E_t(x^t, y^t)$ is defined as the inverse of the corresponding Shephardian distance function.[2]

For all $(a, b) \in \{t, t+1\}^2$, the time-related versions of the Farrell input efficiency measure are given by

$$E^i_a(x^b, y^b) = \inf_{\lambda}\{\lambda : (\lambda x^b, y^b) \in T^a\} \tag{4}$$

if there is some λ such that $(\lambda x^b, y^b) \in T^a$ and $E^i_a(x^b, y^b) = +\infty$ otherwise. Similarly, in the output case, $E^o_a(x^b, y^b) = \sup_{\theta}\{\theta : (x^b, \theta y^b) \in T^a\}$ if there is some θ such that $(x^b, \theta y^b) \in T^a$ and $E^o_a(x^b, y^b) = -\infty$ otherwise.

2.2. Malmquist productivity index

Using the input Farrell measures one can define the input-oriented Malmquist productivity index in base period t as follows:

$$M^i_t(x^t, y^t, x^{t+1}, y^{t+1}) = \frac{E^i_t(x^{t+1}, y^{t+1})}{E^i_t(x^t, y^t)}. \tag{5}$$

Values of this base period t input-oriented Malmquist productivity index above (below) unity reveal productivity growth (decline).

Similarly, a base period $t+1$ input-oriented Malmquist productivity index is defined as follows:

$$M^i_{t+1}(x^t, y^t, x^{t+1}, y^{t+1}) = \frac{E^i_{t+1}(x^{t+1}, y^{t+1})}{E^i_{t+1}(x^t, y^t)}. \tag{6}$$

[2] Farrell (1957, p. 259) in fact verbally defines both input- and output-oriented efficiency measures as being smaller than unity. Under such definition, the input-oriented efficiency measure is the inverse of the input distance function, while the output-oriented efficiency measure simply equals the output distance function.

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

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

$$M^i_{t,t+1} = \sqrt{M^i_t \cdot M^i_{t+1}}, \tag{7}$$

whereby the arguments of the functions are suppressed to save space. Note again that when the geometric mean input-oriented Malmquist productivity index is larger (smaller) than unity, 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 such as the Hicks–Moorsteen productivity index.

2.3. Hicks–Moorsteen productivity index (TFP)

Following the seminal article by Bjurek (1996), a Hicks–Moorsteen productivity (or Malmquist TFP) index with a base period t is defined as the ratio of a Malmquist output quantity index in base period t over a Malmquist input quantity index in the same base period t:

$$HM_t(x^t, y^t, x^{t+1}, y^{t+1}) = \frac{MO_t(x^t, y^t, y^{t+1})}{MI_t(x^t, x^{t+1}, y^t)} \tag{8}$$

whereby the output quantity index is defined as $MO_t(x^t, y^t, y^{t+1}) = \frac{E^o_t(x^t, y^t)}{E^o_t(x^t, y^{t+1})}$ and the input quantity index is defined as $MI_t(x^t, x^{t+1}, y^t) = \frac{E^i_t(x^t, y^t)}{E^i_t(x^{t+1}, y^t)}$. If the Hicks–Moorsteen productivity index is larger (smaller) than unity, then it indicates a gain (loss) in productivity.

Similarly, a base period $t+1$ Hicks–Moorsteen productivity index is defined as follows:

$$HM_{t+1}(x^t, y^t, x^{t+1}, y^{t+1}) = \frac{MO_{t+1}(x^{t+1}, y^{t+1}, y^t)}{MI_{t+1}(x^t, x^{t+1}, y^{t+1})} \tag{9}$$

where we now have for the output quantity index $MO_{t+1}(x^{t+1}, y^{t+1}, y^t) = \frac{E^o_{t+1}(x^{t+1}, y^t)}{E^o_{t+1}(x^{t+1}, y^{t+1})}$ and for the input quantity index $MI_{t+1}(x^t, x^{t+1}, y^{t+1}) = \frac{E^i_{t+1}(x^t, y^{t+1})}{E^i_{t+1}(x^{t+1}, y^{t+1})}$. Again, when the Hicks–Moorsteen productivity index is larger (smaller) than unity, it points to a productivity gain (loss).

To avoid a choice of base year, it is customary to take a geometric mean of these two Hicks–Moorsteen productivity indices (see Bjurek (1996)):

$$HM_{t,t+1} = \sqrt{HM_t \cdot HM_{t+1}}, \tag{10}$$

where arguments of the functions are suppressed for reasons of space. Note once more that when the geometric mean Hicks–Moorsteen productivity index is larger (smaller) than unity, it points to a productivity gain (loss).

A final observation can be made. The denominator of both the Malmquist output and input quantity indices in the base period t Hicks–Moorsteen productivity index compares a "hypothetical" or pseudo-observation consisting of inputs and outputs observed from different periods to a technology in period t. The same remark applies to the numerator for the corresponding Malmquist output and input quantity indices in base period $t+1$. Such "hypothetical" observations do not appear in the Malmquist productivity index, which makes for a somewhat easier interpretation.

2.4. Primal productivity indices: a comparison

We end this section with some remarks regarding the properties of both these primal productivity indices (see also O'Donnell (2012a) for more details).

First, one well-known pitfall of the Malmquist productivity index is that it is not always a TFP index (see Bjurek et al. (1998)). For instance, while its TFP properties are maintained under constant returns to scale, as illustrated by Grifell-Tatjé and Lovell (1995), these properties are not preserved in the presence of variable returns to scale (i.e., a more general technology). By contrast, Bjurek (1996) is the first to state that the Hicks–Moorsteen productivity index has a TFP interpretation (see also Bjurek et al. (1998)). More recently, O'Donnell (2010) shows that profitability change can be decomposed into the product of a Total Factor Productivity (TFP) index and an index measuring relative price changes. While our focus is on TFP as the real part of profitability, O'Donnell (2012b, p. 875–877) elaborates on the potential interactions between both components of profitability change. Following O'Donnell (2012a, 2012b), it is known that many TFP indices can be decomposed into measures of technical change and technical efficiency change (following Nishimizu & Page (1982)), but furthermore into scale efficiency change and mix efficiency change components. Indices that can be decomposed in this way include the Fisher, Törnqvist and Hicks–Moorsteen TFP indices, but not the Malmquist productivity index. In fact, Grosskopf (2003) suggests to call the Malmquist productivity index a technology index. In other words, the Malmquist productivity index just measures local technical change (i.e., the local change of a production frontier allowing for efficiency change), but it cannot be used to measure TFP change in general (in contrast to widespread opinion).[3]

Second, both ratio-based productivity indices can be related to one another under rather stringent conditions. Indeed, one key analytical relation is established in Färe, Grosskopf, and Roos (1996): both indices coincide under constant returns to scale and inverse homotheticity. Some additional relations are mentioned in Bjurek et al. (1998): both indices coincide under constant returns to scale in the case of (i) a single input and multiple outputs, (ii) a single output and multiple inputs, and (iii) when all inputs and/or outputs of a unit change proportionally. Finally, O'Donnell (2012a, p. 258), points out that both indices are equal under constant returns to scale in the absence of technical change.

However, empirical studies comparing both indices are extremely rare. We are aware of only two such studies: Bjurek et al. (1998) report minor differences between both indices, a result confirmed recently in Simões and Marques (2012) where minor deviations are listed between averages, minima and maxima over the short period analyzed.[4] This limited empirical evidence could be taken as a first indication that the above conditions on technology under which both indices coincide do not seem to hold exactly. However, these empirical studies do not formally test for the similarity of both indices. This study fills this void and tests whether their distributions of productivity change are identical under a variety of technology specifications and in the presence of balanced or unbalanced panel data.

[3] Peyrache (in press) defines a radial productivity index (RPI) that makes the explicit connection between the Hicks–Moorsteen TFP index and the Malmquist index and interprets it to reveal the elements missing in the Malmquist index.

[4] While the Malmquist productivity index is very popular, the Hicks–Moorsteen productivity index has so far found rather limited use in applied research. Apart from the above two studies, a to our knowledge rather complete list of empirical applications includes Arjomandi, Harvie, and Valadkhani (2012), Arora and Arora (2012, 2013), Epure, Kerstens, and Prior (2011), Hoang (2011), Nemoto and Goto (2005), O'Donnell (2012a, 2012b), and Zaim (2004).

Third, another problem known since the beginning of this literature is that some of the distance functions constituting the Malmquist productivity index may well be undefined when estimated using general technologies (see Färe et al. (1994), footnote 15). However, empirical studies often ignore reporting on this infeasibility problem. Briec and Kerstens (2009) prove that infeasibilities can occur for an even more general productivity indicator based upon more general distance functions. Thus, even this more general indicator does not satisfy the determinateness property in index theory. By contrast, the Hicks–Moorsteen index satisfies the determinateness axiom, as conjectured by Bjurek (1996) and proven in Briec and Kerstens (2011) under mild conditions (i.e., mainly strong disposability of inputs and outputs).[5]

Fourth, as mentioned in the introduction, both these primal productivity indices can be computed on balanced and unbalanced panel data alike. However, in view of the preceding remark it is critical to distinguish between an infeasibility due to unavailable data (e.g., related to the unbalanced nature of the panel) and a computational infeasibility. The former case could probably better be called a logical impossibility because one simply cannot measure the underlying adjacent period efficiency measures being part of the productivity indices.

Overall, the TFP nature of the Hicks–Moorsteen index and the fact that it can easily be made transitive by a proper choice of basis (underscored by O'Donnell (2012b)) make it undoubtedly deserve greater attention. Transitivity allows for meaningful multi-lateral and multi-temporal comparisons (instead of only binary comparisons). The reader is referred to O'Donnell (2012b) for more details on the economically-relevant axioms a suitable choice of base for the Hicks–Moorsteen index can yield. This assessment echoes the conclusion earlier made by Bjurek et al. (1998) and Lovell (2003) in the same context.

Note that the formulations of the Malmquist and the Hicks–Moorsteen indices above use a variable base year approach (geometric mean). Just like the Hicks–Moorsteen index can be given a fixed base (see O'Donnell (2010)) resulting in attractive properties, also the Malmquist index can be defined with a fixed base (see, e.g., Berg, Førsund, & Jansen (1992)). However, we stick to this choice of variable basis because it is by far the most popular in empirical work employing the Malmquist index (see, e.g., Ouellette & Vierstraete (2004) or Lozano-Vivas & Humphrey (2002)). Note that the choice of basis is not normally going to influence our results fundamentally. We now turn to the treatments for unbalanced panel data found in the literature.

3. Treatments for unbalanced panel data and critiques

One basic strategy found in the empirical literature employing these primal productivity indices consists in making the unbalanced panel somehow balanced. In fact, a variety of strategies can be discerned in the literature.

First, a straightforward strategy consists in simply dropping the observations that are not balanced. One – already cited – example is the article by Hollingsworth and Wildman (2003). Other examples of studies seemingly applying this strategy include Matthews and Zhang (2010) or Sturm and Williams (2004), among others.

Second, sometimes a kind of natural remedy is employed to make the unbalanced panel a balanced one. One example is the backward merger of units: units that merge at some point in time are also treated as merged for the years in the sample preceding the year of the merger. An example of a study adopting this remedy is Tortosa-Ausina, Grifell-Tatjé, Armero, and Conesa (2008).

Third, alternatively some authors resort to a more artificial remedy to make the initially unbalanced panel balanced. One example is the creation of artificial units in an effort to make the panel balanced (see, e.g., Hongliang & Pollitt (2009)). Another example is to somehow complete the missing data. The article by Simões and Marques (2012) assumes that the missing data, which are all situated in the first years of the observation period, are identical to those in the first year with available data.

Other strategies are more elaborate and involve some kind of partial balancing of the data set. For instance, one kind of intermediate solution found in the literature is to balance on a 2-years by 2-years basis. In such a setting, all firms present in each of the adjacent two-year comparison periods (the adjacent-year sample) are maintained (see, e.g., Cummins & Rubio-Misas (2006) for an empirical paper). More in general, one can note that some proposals to average these productivity indices over a variety of base periods are at least partially motivated by the desire to accommodate the case of unbalancedness in panel data (for instance, Asmild & Tam (2007)).

Without claiming to have exhausted the proposals around in the literature, one can state as a preliminary conclusion that most proposals lack general validity (though each and every one of these proposals may be meaningful for certain particular research purposes). Furthermore, these pseudo-solutions simply beg the question about the impact of unbalancedness on productivity measurement.

It is well-known that unbalancedness can occur due to delayed entry, early exit, or intermittent nonresponse. Another important distinction is that the lack of balance can be either planned (designed) as, for instance, in the case of rotating panels, or unplanned. In the latter unplanned case, non-responses are called missing data and these represent a potential source of bias. This is in particular the case in situations in which the mechanisms for missingness are related to the phenomenon being modeled (i.e., attrition bias). See Baltagi and Song (2006) or Frees (2004) for more details.

In the context of productivity measurement, attrition bias is a known issue (Van Beveren (2010) offers a survey of estimation issues) and it has regularly been reported in some parts of the literature (see, e.g., Drake & Simper (2002) or Foster, Haltiwanger, & Syverson (2008) for some recent examples). However, we are aware of only few articles discussing these issues in the efficiency and productivity literature using frontier specifications. For instance, Byrnes (1991) is the only study we are aware of explicitly analyzing selectivity bias related to public or private ownership of water utilities in a parametric cost efficiency frontier context. In a somewhat similar vein, Scully (1994) and Wheelock and Wilson (2000) use survival analysis where efficiency scores are part of the explanatory variables to model sport coach survival and bank failures or acquisitions respectively. We are unaware of any article reporting attrition bias while employing the frontier-based primal productivity indices analyzed in this study.

However, unbalancedness is in practice an unknown mix of unplanned and planned elements. Furthermore, the exact reason for the missing data (i.e., delayed entry, early exit, or intermittent nonresponse) is rarely known to the empirical researcher. If the exact reason for the missing data is known to the analyst, then it seems obvious that one should exploit this knowledge to measure the contribution of entering and exiting firms to productivity growth (see Griliches & Regev (1995) or more recently Diewert & Fox (2010)).

In general, it would seem useful to at least document the eventual impact of unbalancedness vs. balancedness in these primal productivity measures. Only when the impact would be negligible,

[5] Zaim (2004) employs a Hicks–Moorsteen index to measure environmental performance imposing weak disposal in the bad outputs that are jointly produced with the good outputs. Not entirely surprisingly, he reports some infeasibilities for this Hicks–Moorsteen environmental performance index.

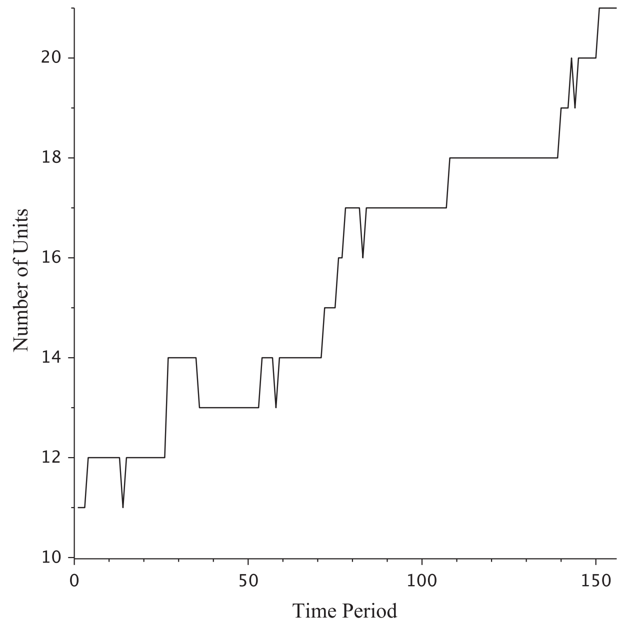

Fig. 1. Available observations per period for the Atkinson data.

one could envision ignoring the issue. In the next section, we turn to this empirical exercise.

4. Data, methodology, and empirical illustration

In this section, we first present the samples used for the empirical illustration. Then, we present the various technologies employed to compute the efficiency measures underlying both the Malmquist and Hicks–Moorsteen productivity indices. Thereafter, we provide the empirical results.

4.1. Data description

For this empirical part, we use two publicly available data bases from the Journal of Applied Econometrics Data Archive that happen to analyze unbalanced panel data.[6] For our purpose, we balance these unbalanced panels to illustrate the effects on both primal productivity indices.

The first data base is a rather short unbalanced panel of three years (1984–1986) of French fruit producers based on annual accounting data collected in a survey (see Ivaldi, Ladoux, Ossard, & Simioni (1996) for details). Farms are selected on mainly two criteria: (i) the production of apples must be positive, and (ii) the acreage of the orchard must be at least five acres. As a description of technology, three aggregate inputs produce two aggregate outputs. The three inputs are: (i) capital (including land), (ii) labor, and (iii) materials. The two aggregate outputs are (i) the production of apples, and (ii) an aggregate of alternative products.

In total, 184 farms are available in the data base of which 130, 135 and 140 have records in 1984, 1985 and 1986, respectively. Thus, the unbalanced panel contains 405 observations in total. The balanced panel, containing only those farms for which records are available for all years, consists of only 92 farms. This yields an overall total of 276 observations. Thus, imposing balancedness amounts to eliminating about 32% of the information in the sample. Further summary statistics for all observations and details on the definitions of all variables are available in Appendix 2 in Ivaldi et al. (1996).

The second data base is a small sample with a single output analyzed by Atkinson and Dorfman (2009). This panel data set contains 16 hydro-electric power plants in Chile observed on a monthly basis from April 1986 to December 1997. These plants generate a single output (i.e., electricity generated). Also the quantities of three inputs are available (i.e., labor, capital, and water). Except for capital, all remaining flow variables are expressed in physical units. Basic descriptive statistics for the three input quantities as well as for the single output as well as more details on these data are available in Atkinson and Dorfman (2009).

The unbalanced panel contains 1935 monthly observations in total. The number of observations per plant range from 49 to 141. Fig. 1 plots the number of observations per time period. Two striking observations can be made. First, the trend of the number of observations moves upwards. Second, there is quite a bit of variation around this trend. Note also that there are 4 observations with a zero output, while 1 observation has a sequence of 18 months with no output at all.

4.2. Specifications of technologies for the efficiency computations

For the empirical application, we employ a variety of non-parametric technologies. In particular, we use both convex and non-convex technologies and both constant and variable returns to scale assumptions. The reason for this choice is that in these non-parametric technologies the potential infeasibilities of the estimated distance functions and the resulting indeterminateness of the productivity index is at least well understood (see Briec & Kerstens (2009)). Otherwise, we do not enter into the debates which technology specification or estimation method is better when computing productivity indices in the potential presence of measurement error, inefficiencies, technology misspecification, etc.[7]

Let K be the number of units. A unified algebraic presentation for a technology satisfying some combination of the above axioms is:

$$T^{\Lambda,\Gamma} = \left\{ (x,y) \in \mathbb{R}_+^{n+p} : y_i \leqslant \sum_{k=1}^{K} \delta z_k y_{ki},\ (i = 1,\dots,p), \right.$$
$$\left. \sum_{k=1}^{K} \delta z_k x_{kj} \leqslant x_n,\ (j = 1,\dots,n), z \in \Lambda, \delta \in \Gamma \right\},$$

where $\Lambda \in \{C, NC\}$, with $C = \{z \in \mathbb{R}_+^K : \sum_{k=1}^{K} z_k = 1\}$ and $NC = \{z \in \mathbb{R}_+^K : \sum_{k=1}^{K} z_k = 1 \text{ and } \forall k = 1,\dots,K : z_k \in \{0,1\}\}$, and where $\Gamma \in \{CRS, VRS\}$, with $CRS = \mathbb{R}_+$ and $VRS = \{1\}$.

From activity analysis, z is the vector of activity variables that indicates the intensity at which a particular activity is employed in constructing the reference technology by forming convex or non-convex combinations of observations constituting the best practice frontier (see Briec, Kerstens, & Vanden Eeckaut (2004)).

Axioms (T.1)–(T.4) are maintained in the non-convex case, while the convex case also imposes (T.5). In addition, both these technologies can impose constant returns to scale (T.6) (abbreviated CRS) rather than flexible returns to scale (abbreviated VRS). This unified specification is non-linear, but it can be straightforwardly linearized in the convex case. For the non-convex case, it basically involves solving either some non-linear mixed integer programs, or some scaled vector dominance algorithms.

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

[7] See Van Biesebroeck (2007) or Giraleas, Emrouznejad, and Thanassoulis (2012) for a more complete discussion and Monte-Carlo evidence.

Table 1
Descriptive statistics for Malmquist and Hicks–Moorsteen productivity indices for the Ivaldi data.

	Malmquist				Hicks–Moorsteen			
	Unbalanced		Balanced		Unbalanced		Balanced	
	1984–1985	1985–1986	1984–1985	1985–1986	1984–1985	1985–1986	1984–1985	1985–1986
$T^{C,CRS}$								
n	110	111	92	92	110	111	92	92
Average	1.1804	1.0964	1.1973	1.1089	1.1793	1.0965	1.1934	1.1070
Stand. Dev.	1.1605	0.7610	1.2354	0.8032	1.1556	0.7540	1.2234	0.7938
Min	0.3357	0.1900	0.3244	0.1910	0.3418	0.1913	0.3264	0.1919
Max	11.7008	6.9671	11.7158	6.9671	11.6472	6.8672	11.6385	6.8672
# Contrad. Res. Mq./HM	0	0	0	1				
$T^{C,VRS}$								
n	107	108	89	89	110	111	92	92
Average	0.9348	1.1288	0.9529	1.1416	1.1500	1.1675	1.1611	1.1812
Stand. Dev.	0.2540	0.2825	0.2643	0.2889	1.1368	0.7237	1.1999	0.7772
Min	0.5245	0.3832	0.5125	0.3978	0.3318	0.1893	0.3472	0.1925
Max	1.8934	1.7912	1.8938	1.8123	11.5799	6.4165	11.5746	6.4861
# Contrad. Res. Mq./HM	21	26	15	20				
$T^{NC,CRS}$								
n	110	111	92	92	110	111	92	92
Average	1.1278	1.1087	1.1310	1.1100	1.1300	1.1003	1.1359	1.0955
Stand. Dev.	0.7936	0.8085	0.8113	0.8347	0.8003	0.7421	0.8191	0.7836
Min	0.2284	0.2846	0.2284	0.2793	0.2507	0.2625	0.2507	0.2829
Max	6.7987	7.9900	7.0040	7.6651	6.7987	7.0934	7.0040	7.0399
# Contrad. Res. Mq./HM	8	4	6	4				
$T^{NC,VRS}$								
n	105	107	87	87	110	111	92	92
Average	0.9774	1.0877	0.9947	1.0947	1.0992	1.1402	1.0995	1.1215
Stand. Dev.	0.2949	0.3211	0.3394	0.3485	0.6649	0.6579	0.6649	0.6714
Min	0.3501	0.4957	0.3501	0.4228	0.4015	0.1668	0.4022	0.1668
Max	2.1497	1.9756	2.3750	2.0105	5.3857	5.7232	5.2018	5.7377
# Contrad. Res. Mq./HM	26	12	18	16				

4.3. Empirical results for the primal productivity indices

Table 1 contains basic descriptive statistics for both the Malmquist and Hicks–Moorsteen productivity indices with the balanced and unbalanced panel data and using several technologies for the French fruit producers. This table is structured as follows: (i) The first four columns list the Malmquist, the last four columns report the Hicks–Moorsteen results. (ii) Within the latter distinction, the first two columns always contain the results for the unbalanced panel, while columns three and four each time display the balanced panel results. (iii) Horizontally, we first distinguish between convex and non-convex technologies. (iv) Then, we separately report both CRS and VRS assumptions imposed on a given technology.

Computation of these descriptive statistics is performed over the productivity indices available. To give an example, the Malmquist index for the unbalanced panel results in valid results for 110 farms for the period 1984–1985. Consequently, all corresponding descriptive statistics are computed for these 110 valid results. Obviously, due to a priori removal of data, the number of valid results is 92 for the balanced panel, unless computational infeasibilities occur. For example, in case of the specification $T^{C,VRS}$ only 89 valid computations are recorded for the periods 1984–1985 and 1985–1986 because of 3 such computational infeasibilities in each of these periods. Just to provide some further interpretation, for the unbalanced sample under $T^{C,CRS}$ specification productivity in 1985 was about 18.04% higher than in 1984 according to the Malmquist index, while the Hicks–Moorsteen index indicates a productivity growth of about 17.93% over the same period. By contrast, for the unbalanced sample under $T^{NC,VRS}$ the Malmquist productivity index reports a 2.26% decline of 1985 compared to 1984 according while the Hicks–Moorsteen index points to a productivity growth of about 9.92% over the same period.

Several conclusions jump out. First, Malmquist and Hicks–Moorsteen productivity indices sometimes seem to disagree on the nature of productivity change at the sample level: cases in point are the specifications $T^{C,VRS}$ and $T^{NC,VRS}$ for the years 1984–1985. Second, the descriptive statistics for both indices are rather different when comparing the balanced and the unbalanced cases. It is an open question whether any of these differences is statistically significant (see infra). Third, these descriptive statistics seem rather robust across the several specifications of technology, with the exception of the specifications $T^{C,VRS}$ and $T^{NC,VRS}$.

Obviously, such sample level results may potentially hide a lot of conflicting results at the level of individual observations. Therefore, the last horizontal line for each specification reports the number of observations for which Malmquist and Hicks–Moorsteen indices offer contradictory results: one pointing to productivity decline while the other shows productivity growth, or the other way around (denoted "# Contrad. Res. Mq./HM"). Two conclusions arise. First, under CRS there is just a single case of opposite results under convexity, while several cases arise under non-convexity. Second, quite a bit of contradictory results arise under the VRS hypothesis independent of the convexity axiom.

Table 2 relates to the Chilean hydro-electric power plants and is similar in structure to Table 1, except that only overall results are reported. Given the numbers of periods involved, it is simply impossible to report period to period results. We can now state the following conclusions. First, Malmquist and Hicks–Moorsteen productivity indices do not seem to disagree on the nature of productivity change at the sample level. Distributions seem to be close to one another especially for the CRS technologies. Second, descriptive statistics for both indices are again different when comparing the balanced and the unbalanced samples. Third, these descriptive statistics are more robust for CRS than for VRS specifications. Again, these sample level results may well hide conflicting results

Table 2
Descriptive statistics for Malmquist and Hicks–Moorsteen productivity indices for the Atkinson data.

	Malmquist		Hicks–Moorsteen	
	Unbalanced	Balanced	Unbalanced	Balanced
$T^{C,CRS}$				
n	2412	1085	2412	1085
Average	1.0436	1.0181	1.0437	1.0181
Stand. Dev.	0.3493	0.1912	0.3498	0.1912
Min	0.1687	0.3970	0.1687	0.3970
Max	6.1295	2.4833	6.1295	2.4833
# Contrad. Res. Mq./HM	0	12		
$T^{C,VRS}$				
n	2227	931	2412	1085
Average	1.0397	1.0247	1.0496	1.0254
Stand. Dev.	0.3084	0.2350	0.4034	0.2484
Min	0.1853	0.1859	0.1040	0.1774
Max	5.0747	2.6642	8.3617	4.6910
# Contrad. Res. Mq./HM	213	69		
$T^{NC,CRS}$				
n	2412	1085	2412	1085
Average	1.0436	1.0177	1.0436	1.0177
Stand. Dev.	0.3628	0.1966	0.3628	0.1966
Min	0.1660	0.3970	0.1660	0.3970
Max	8.0005	2.6449	8.0005	2.6449
# Contrad. Res. Mq./HM	11	18		
$T^{NC,VRS}$				
n	2227	933	2412	1085
Average	1.1555	1.1681	1.0509	1.0396
Stand. Dev.	0.7087	0.7161	0.4054	0.3093
Min	0.1510	0.2793	0.0857	0.1941
Max	6.4895	3.4799	10.2198	4.6910
# Contrad. Res. Mq./HM	313	107		

for individual observations. The same two conclusions as before can be deduced. However, in this sample the amount of contradictory results seems somewhat larger under the non-convexity axiom.

Table 3 reports on the relative presence of infeasibilities due to unavailable data (denoted ‘‘na’’) and the computational infeasibilities (denoted ‘‘Inf’’) for the fruit producers. The first line lists the unavailable data for both unbalanced and balanced panel data and is common to both the Malmquist and Hicks–Moorsteen indices. The second part of the table reports the infeasibilities in the Malmquist index depending on the various technology specifications.

Three conclusions emerge from studying this table. First, infeasibilities due to unavailable data amount to 50% in the balanced case, while these vary around 40% depending on the exact year in the unbalanced case. This amounts to a gain of about 10% in the amount of information included in the estimates. Second, despite this gain in the amount of information, the percentage of computational infeasibilities seems rather stable when comparing the balanced and the unbalanced cases. For the Malmquist index, the computational infeasibilities vary between 0.00% and 2.72% in both the unbalanced and the balanced cases depending on the technology specification. In the $T^{C,VRS}$ specification, the amount of computational infeasibilities remains stable at 1.63% for both periods and in both the balanced and unbalanced cases. No computational infeasibilities occur for the Malmquist index with the $T^{C,CRS}$ and $T^{NC,CRS}$ specifications. Third, the Hicks–Moorsteen index does not have a single computational infeasibility for all the technology and panel specifications over all periods. This is why it is not reported in Table 3.

Table 4
Productivity indices under various specifications for the Atkinson data: non-availabilities (‘‘na’’) and infeasibilities (‘‘Inf’’).

		Unbalanced	Balanced
	% na	25.90	66.67
Malmquist			
$T^{C,CRS}$	% Inf	0.00	0.00
$T^{C,VRS}$	% Inf	5.68	4.73
$T^{NC,CRS}$	% Inf	0.00	0.00
$T^{NC,VRS}$	% Inf	5.68	4.67

Table 5
Li-test results for the Ivaldi data under various specifications.

	Unbalanced vs. balanced		Malmquist vs. Hicks–Moorsteen	
	Malmquist	Hicks–Moorsteen	Unbalanced	Balanced
$T^{C,CRS}$	−1.2345	−1.2279	−0.0064	−0.0029
$T^{C,VRS}$	−1.1133	−1.3342	4.0058	2.8723
$T^{NC,CRS}$	−1.2718	−1.1736	0.0357	−0.0900
$T^{NC,VRS}$	−1.0169	−1.1432	3.1633	0.7237

Li test: critical values at 1% level = 2.33 (***); 5% level = 1.64 (**); 10% level = 1.28 (*).

Table 4 focuses on the power plants and has a structure identical to Table 3. Its analysis again yields three key conclusions. First, infeasibilities due to unavailable data (‘‘na’’) vary between 25.90% and 66.67% in the unbalanced and the balanced case respectively. This implies a gain in the amount of information included in the estimates that reaches a staggering 40%. Second, the percentage of computational infeasibilities (‘‘Inf’’) is higher now compared to the first data set, but it remains quite stable when comparing the balanced and the unbalanced cases (it varies between 4.67% and 5.68% depending on the technology specification). Again, for the Malmquist index with the $T^{C,CRS}$ and $T^{NC,CRS}$ specifications we can detect no computational infeasibilities. Third, the Hicks–Moorsteen index is always feasible.

Table 5 formally tests for the differences between the densities of these productivity indices with a test statistic proposed by Li (1996) (see also Fan & Ullah (1999) for a refinement) that is valid for both dependent and independent variables for the French fruit producers. Note that dependency is a characteristic for these frontier estimators (e.g., efficiency levels depend, among others, on sample size). The null hypothesis states the equality of both balanced and unbalanced distributions for a given productivity index and underlying specification of technology. The first two columns test between balanced and unbalanced samples for a given

Table 3
Productivity indices under various specifications for the Ivaldi data: non-availabilities (‘‘na’’) and infeasibilities (‘‘Inf’’).

		Unbalanced			Balanced		
		1984–1985	1985–1986	Overall	1984–1985	1985–1986	Overall
	% na	40.22	39.67	39.95	50.00	50.00	50.00
Malmquist							
$T^{C,CRS}$	% Inf	0.00	0.00	0.00	0.00	0.00	0.00
$T^{C,VRS}$	% Inf	1.63	1.63	1.63	1.63	1.63	1.63
$T^{NC,CRS}$	% Inf	0.00	0.00	0.00	0.00	0.00	0.00
$T^{NC,VRS}$	% Inf	2.72	2.17	2.45	2.72	2.72	2.72

Table 6
Li-test results for the Atkinson data under various specifications.

	Unbalanced vs. balanced		Malmquist vs. Hicks–Moorsteen	
	Malmquist	Hicks–Moorsteen	Unbalanced	Balanced
$T^{C,CRS}$	5.5892	5.9827	0.0034	0.0002
$T^{C,VRS}$	0.8199	5.1901	9.4334	1.9531
$T^{NC,CRS}$	4.0646	3.9623	0.0000	0.0000
$T^{NC,VRS}$	15.8214	2.0118	247.8031	177.4574

Li test: critical values at 1% level = 2.33 (***); 5% level = 1.64 (**); 10% level = 1.28 (*).

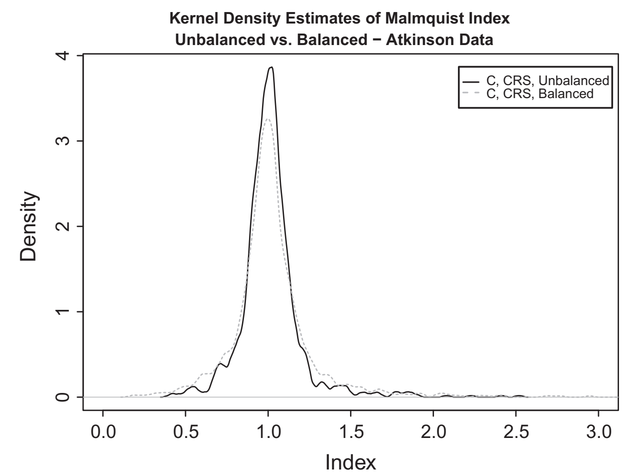

Fig. 2. Kernel density of balanced and unbalanced Malmquist indices under $T^{C,CRS}$ for Atkinson data.

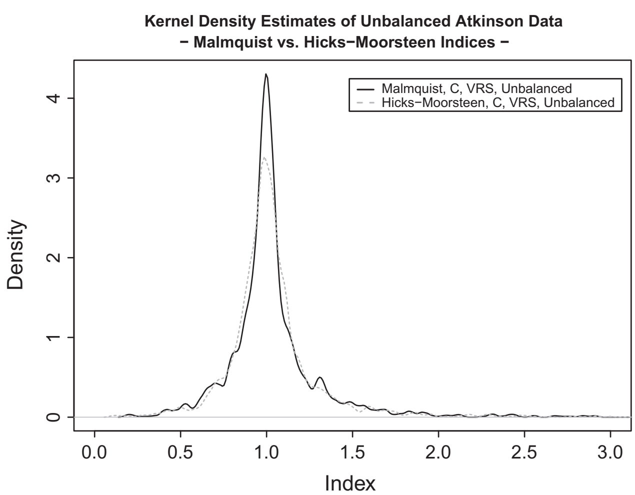

Fig. 3. Kernel density of unbalanced Malmquist and Hicks–Moorsteen indices under $T^{C,VRS}$ for Atkinson data.

productivity index. The last two columns test between Malmquist and Hicks–Moorsteen productivity indices for either unbalanced or balanced data sets.

For the Ivaldi data, the differences in densities between both balanced and unbalanced data sets turn out to be non-significant for this sample. However, the differences between Malmquist and Hicks–Moorsteen productivity indices are significant for most *VRS* specifications at 10% significance levels or more, except for the balanced case. For the *CRS* specifications Malmquist and Hicks–Moorsteen productivity indices have no statistically different densities (though in one case the 10% significance level is rather close).

Table 6 is identical in structure but relates to the power plant sample. Now the differences in densities between both balanced and unbalanced panel data sets are always significant for the Atkinson data, except for the Malmquist index under the $T^{C,VRS}$ specification. Once more, the differences between Malmquist and Hicks–Moorsteen productivity indices are significantly different at varying significance levels for all *VRS* specifications, but are clearly not different at all in the *CRS* case.

To appreciate the observed differences in more detail we also plot kernel densities for a selection of productivity indices for a variety of frontier specifications for the Atkinson data. Fig. 2 plots the densities for the Malmquist index for all years under $T^{C,CRS}$ for the balanced vs. the unbalanced samples. For the same years and the unbalanced sample, Fig. 3 plots the densities of Malmquist vs. Hicks–Moorsteen index under $T^{C,VRS}$. Note that to facilitate comparison, the densities on each figure are estimated with a common bandwidth. In general, these densities illustrate two of the statistically significant differences already revealed by the Li-test above in Table 6.

5. Conclusions

Using two data bases (French fruit producers and Chilean hydro-electric power plants), this contribution is – to the best of our knowledge – the first to empirically illustrate and formally test for the differences between using (i) either unbalanced or balanced panel data when computing frontier estimates for two frontier-based primal productivity indices, and (ii) either the Malmquist or the Hicks–Moorsteen productivity index using both unbalanced and balanced panel data.

In particular, the main empirical results regarding the effect of balancing an unbalanced panel data is that in the balanced case one can loose substantial amounts of information (around 10–40% in our samples). The differences between the primal productivity indices computed relative to unbalanced or balanced panel data can be significantly different depending on the data set studied. One may conjecture that this probably depends on the exact nature of the unbalancedness. This certainly needs further investigation. Having documented the non-negligible impact of balancedness on these primal productivity measures, it is no longer an option to ignore this issue.

As to the question whether the Malmquist and Hicks–Moorsteen indices are empirically indistinguishable or not, the differences between both primal productivity indices turn out to be significantly different for all flexible returns to scale technology specifications in both data sets. Notice that these test at the sample level may hide large differences for individual observations. For these two samples, the Malmquist productivity index maintains a TFP interpretation by approximation only when measured relative to a constant returns to scale technology. Obviously, the robustness of these findings is not guaranteed. This particular issue definitively warrants further testing. In case one wants to use a Malmquist index, it may be prudent to test for constant returns to scale and inverse homotheticity (see Cavaignac & Briec (2007) for the latter test).[8] Whether such tests are useful exercises or not given longstanding misgivings on homothetic structures in production theory is left to the reader.[9] If one wants to be on the safe side, then one conclusion is that in case the interest centers on TFP measurement it is probably wise to immediately opt for the Hicks–Moorsteen index.

[8] For statistical inference on both tests: see Simar and Wilson (2002, 2001) respectively.

[9] Already Samuelson and Swamy (1974, p. 592), concluded: "Empirical experience is abundant that the Santa Claus hypothesis of homotheticity in tastes and in technical change is quite unrealistic."

Obviously, there remain open challenges for future research. One obvious extension is to duplicate this research by comparing the difference- instead of ratio-based Luenberger indicator and its TFP counterpart the Luenberger–Hicks–Moorsteen indicator (see Briec & Kerstens (2004)). Though still less popular than their ratio-based counterpart indices, these indicators have recently found their way in several empirical studies (e.g., Barros, de Menezes, Peypoch, Solonandrasana, & Cabral Vieira (2008)). While inferential issues have been extensively studied when using parametric technology specifications estimated using unbalanced panel data, it remains somehow an open issue in the case of non-parametric specifications as employed in this study. When using unbalanced data, a key benefit is a larger sample. However, the technology per year depends on varying numbers of observations such that the precision of the estimates varies over the years. When balanced data is used, the drawback is a smaller sample, but at least the precision of the estimates does not vary over the years.[10]

[10] In fact, this problem has explicitly motivated some authors (e.g., Odeck (2008)) to estimate these primal productivity indices using balanced data even when an unbalanced panel were available.

References

Arjomandi, A., Harvie, C., & Valadkhani, A. (2012). An empirical analysis of Irans banking performance. *Studies in Economics and Finance, 29*(4), 287–300.
Arora, H., & Arora, P. (2012). Bank productivity measurement using Hicks–Moorsteen indices: Evidence from Indian public sector banks. *International Journal of Business Performance Management, 13*(3), 386–407.
Arora, H., & Arora, P. (2013). Measuring and decomposing productivity change using Hicks–Moorsteen index numbers: Evidence from Indian banks. *International Journal of Productivity and Quality Management, 11*(1), 73–95.
Asmild, M., & Tam, F. (2007). Estimating global frontier shifts and global Malmquist indices. *Journal of Productivity Analysis, 27*(2), 137–148.
Atkinson, S., Cornwell, C., & Honerkamp, O. (2003). Measuring and decomposing productivity change: Stochastic distance function estimation versus data envelopment analysis. *Journal of Business and Economic Statistics, 21*(2), 284–294.
Atkinson, S., & Dorfman, J. (2009). Feasible estimation of firm-specific allocative inefficiency through bayesian numerical methods. *Journal of Applied Econometrics, 24*(4), 675–697.
Baltagi, B., & Song, S. (2006). Unbalanced panel data: A survey. *Statistical Papers, 47*(4), 493–523.
Barros, C., de Menezes, A., Peypoch, N., Solonandrasana, B., & Cabral Vieira, J. (2008). An analysis of hospital efficiency and productivity growth using the Luenberger indicator. *Health Care Management Science, 11*(4), 373–381.
Berg, S., Førsund, F., & Jansen, E. (1992). Malmquist indices of productivity growth during the deregulation of Norvegian banking, 1980–1989. *Scandinavian Journal of Economics, 94*(S), 211–228.
Bjurek, H. (1996). The Malmquist total factor productivity index. *Scandinavian Journal of Economics, 98*(2), 303–313.
Bjurek, H., Førsund, F., & Hjalmarsson, L. (1998). Malmquist productivity indices: An empirical investigation. In R. Färe, S. Grosskopf, & R. Russell (Eds.), *Index numbers: Essays in honour of Sten Malmquist* (pp. 217–239). Boston: Kluwer.
Briec, W., & Kerstens, K. (2004). A Luenberger-Hicks–Moorsteen productivity indicator: Its relation to the Hicks–Moorsteen productivity index and the Luenberger productivity indicator. *Economic Theory, 23*(4), 925–939.
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.
Briec, W., Kerstens, K., & Vanden Eeckaut, P. (2004). Non-convex technologies and cost functions: Definitions, duality and nonparametric tests of convexity. *Journal of Economics, 81*(2), 155–192.
Byrnes, P. (1991). Estimation of cost frontiers in the presence of selectivity bias: Ownership and efficiency of water utilities. In G. Rhodes (Ed.). *Advances in econometrics* (vol. 9, pp. 121–137). Greenwich: JAI.
Cavaignac, L., & Briec, W. (2007). Comment: Testing for inverse homotheticity: A nonparametric approach. *Japanese Economic Review, 58*(4), 524–531.
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.
Coelli, T. (1996). *A guide to DEAP version 2.1: A data envelopment analysis (computer) program*. Discussion Paper CEPA WP 96/08, Centre for Efficiency and Productivity Analysis (CEPA), University of Queensland, School of Economics.
Cummins, J., & Rubio-Misas, M. (2006). Deregulation, consolidation, and efficiency: Evidence from the Spanish insurance industry. *Journal of Money Credit and Banking, 38*(2), 323–356.
Diewert, W., & Fox, K. (2010). On measuring the contribution of entering and exiting firms to aggregate productivity growth. In W. Diewert, B. Balk, D. Fixler, K. Fox, & A. Nakamura (Eds.). *Price and productivity measurement: Index number theory* (Vol. 6, pp. 41–66). Bloomington: Trafford Press.
Drake, L., & Simper, R. (2002). Economies of scale in UK building societies: A re-appraisal using an entry/exit model. *Journal of Banking and Finance, 26*(12), 2365–2382.
Epure, M., Kerstens, K., & Prior, D. (2011). Technology-based total factor productivity and benchmarking: New proposals and an application. *Omega, 39*(6), 608–619.
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., 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.
Farrell, M. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society Series A: General, 120*(3), 253–281.
Foster, L., Haltiwanger, J., & Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review, 98*(1), 394–425.
Frees, E. (2004). *Longitudinal and panel data: Analysis and applications in the social sciences*. Cambridge University Press.
Giraleas, D., Emrouznejad, A., & Thanassoulis, E. (2012). Productivity change using growth accounting and frontier-based approaches- Evidence from a Monte Carlo analysis. *European Journal of Operational Research, 222*(3), 673–683.
Grifell-Tatjé, E., & Lovell, C. (1995). A note on the Malmquist productivity index. *Economics Letters, 47*(2), 169–175.
Griliches, Z., & Regev, H. (1995). Firm productivity in Israeli industry 1979–1988. *Journal of Econometrics, 65*(1), 175–203.
Grosskopf, S. (2003). Some remarks on productivity and its decompositions. *Journal of Productivity Analysis, 20*(3), 459–474.
Hoang, V.-N. (2011). Measuring and decomposing changes in agricultural productivity, nitrogen use efficiency and cumulative exergy efficiency: Application to OECD agriculture. *Ecological Modelling, 222*(1), 164–175.
Hollingsworth, B., & Wildman, J. (2003). The efficiency of health production: Re-estimating the WHO panel data using parametric and non-parametric approaches to provide additional information. *Health Economics, 12*(6), 493–504.
Hongliang, Y., & Pollitt, M. (2009). Incorporating both undesirable outputs and uncontrollable variables into DEA: The performance of chinese coal-fired power plants. *European Journal of Operational Research, 197*(3), 1095–1105.
Hulten, C. (2001). Total factor productivity: A short biography. In C. Hulten, E. Dean, & M. Harper (Eds.), *New developments in productivity analysis* (pp. 1–47). Chicago: University of Chicago Press.
Ivaldi, M., Ladoux, N., Ossard, H., & Simioni, M. (1996). Comparing Fourier and Translog specifications of multiproduct technology: Evidence from an incomplete panel of French farmers. *Journal of Applied Econometrics, 11*(6), 649–668.
Li, Q. (1996). Nonparametric testing of closeness between two unknown distribution functions. *Econometric Reviews, 15*(1), 261–274.
Lovell, C. (2003). The decomposition of Malmquist productivity indexes. *Journal of Productivity Analysis, 20*(3), 437–458.
Lozano-Vivas, A., & Humphrey, D. (2002). Bias in Malmquist index and cost function productivity measurement in banking. *International Journal of Production Economics, 76*(2), 177–188.
Matthews, K., & Zhang, N. X. (2010). Bank productivity in China 1997–2007: Measurement and convergence. *China Economic Review, 21*(4), 617–628.
Nemoto, J., & Goto, M. (2005). Productivity, efficiency, scale economies and technical change: A new decomposition analysis of TFP applied to the Japanese prefectures. *Journal of the Japanese and International Economies, 19*(4), 617–634.
Nishimizu, M., & Page, J. (1982). Total factor productivity growth, technological progress and technical efficiency change: Dimensions of productivity change in Yugoslavia, 1965–1978. *Economic Journal, 92*(368), 920–936.
Odeck, J. (2008). How efficient and productive are road toll companies? Evidence from Norway. *Transport Policy, 15*(4), 232–241.
O'Donnell, C. (2010). Measuring and decomposing agricultural productivity and profitability change. *Australian Journal of Agricultural and Resource Economics, 54*(4), 527–560.
O'Donnell, C. (2012a). An aggregate quantity-price framework for measuring and decomposing productivity and profitability change. *Journal of Productivity Analysis, 38*(3), 255–272.
O'Donnell, C. (2012b). Nonparametric estimates of the components of productivity and profitability change in U.S. agriculture. *American Journal of Agricultural Economics, 94*(4), 873–890.
Ouellette, P., & Vierstraete, V. (2004). Technological change and efficiency in the presence of quasi-fixed inputs: A DEA application to the hospital sector. *European Journal of Operational Research, 154*(3), 755–763.
Peyrache, A. (in press). Hicks–Moorsteen versus Malmquist: A connection by means of a radial productivity index. *Journal of Productivity Analysis*. http://dx.doi.org/10.1007/s11123-013-0350-2.

Samuelson, P. A., & Swamy, S. (1974). Invariant economic index numbers and canonical duality: Survey and synthesis. *American Economic Review, 64*(4), 566–593.

Scully, G. (1994). Managerial efficiency and survivability in professional team sports. *Managerial and Decision Economics, 15*(5), 403–411.

Simões, P., & Marques, R. (2012). Influence of regulation on the productivity of waste utilities. What can we learn with the portuguese experience? *Waste Management, 32*(6), 1266–1275.

Simar, L., & Wilson, P. (2001). Testing restrictions in nonparametric efficiency models. *Communications in Statistics: Simulation and Computation, 30*(1), 159–184.

Simar, L., & Wilson, P. (2002). Non-parametric tests of returns to scale. *European Journal of Operational Research, 139*(1), 115–132.

Sturm, J.-E., & Williams, B. (2004). Foreign bank entry, deregulation and bank efficiency: Lessons from the Australian experience. *Journal of Banking and Finance, 28*(7), 1775–1799.

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

Tsekouras, K., Pantzios, C., & Karagiannis, G. (2004). Malmquist productivity index estimation with zero-value variables: The case of greek prefectural training councils. *International Journal of Production Economics, 89*(1), 95–106.

Van Beveren, I. (2010). Total factor productivity estimation: A practical review. *Journal of Economic Surveys, 26*(1), 98–128.

Van Biesebroeck, J. (2007). Robustness of productivity estimates. *Journal of Industrial Economics, 55*(3), 529–569.

Wheelock, D., & Wilson, P. (2000). Why do banks disappear? The determinants of U.S. bank failures and acquisitions. *Review of Economics and Statistics, 82*(1), 127–138.

Zaim, O. (2004). Measuring environmental performance of state manufacturing through changes in pollution intensities: A DEA framework. *Ecological Economics, 48*(1), 37–47.