Working Paper Series n° 634

June, 2009

"Forecasting the Malmquist Productivity Index"

Alexandra DASKOVSKA, Léopold SIMAR and Sébastien VAN BELLEGEM

Institut d'Economie Industrielle (IDEI) – Manufacture des Tabacs
Aile Jean-Jacques Laffont – 21, allée de Brienne – 31000 TOULOUSE – FRANCE
Tél. + 33(0)5 61 12 85 89 – Fax + 33(0)5 61 12 86 37 – **www.idei.fr** – contact@cict.fr

INSTITUT D'ECONOMIE INDUSTRIELLE

Forecasting the Malmquist Productivity Index *

Alexandra Daskovska [1] Léopold Simar [2] Sébastien Van Bellegem [3]

This version: June, 11, 2009

Abstract

The Malmquist Productivity Index (MPI) suggests a convenient way of measuring the productivity change of a given unit between two consequent time periods. Until now, only a static approach for analyzing the MPI was available in the literature. However, this hides a potentially valuable information given by the evolution of productivity over time. In this paper, we introduce a dynamic procedure for forecasting the MPI. We compare several approaches and give credit to a method based on the assumption of circularity. Because the MPI is not circular, we present a new decomposition of the MPI, in which the time-varying indices are circular. Based on that decomposition, a new working dynamic forecasting procedure is proposed and illustrated. To construct prediction intervals of the MPI, we extend the bootstrap method in order to take into account potential serial correlation in the data. We illustrate all the new techniques described above by forecasting the productivityt index of 17 OCDE countries, constructed from their GDP, labor and capital stock.

Keywords: Malmquist productivity index, circularity, efficiency, smooth bootstrap, forecasting intervals

*This work was supported by the "Agence National de le Recherche" under contract ANR-09-JCJC-0124-01, by the contract "Projet d'Actions de Recherche Concertées" nr 07/12-002 of the "Communauté franaise de Belgique" granted by the "Académie universitaire Louvain", and the IAP research network nr P6/03 of the Belgian Government (Belgian Science Policy)

[1]Université catholique de Louvain (Institut de statistique)

[2]Université catholique de Louvain (Institut de statistique)

[3]Toulouse School of Economics (GREMAQ)

1 Introduction

The Malmquist Productivity Index (MPI) is a bilateral index that can be used to compare the production technology of two economies. It was introduced by Caves, Christensen and Diewert (1982) who named it after Malmquist (1953), who proposed to construct quantity indices as ratios of distance functions for use in consumption analysis. Although it was initially developed in the context of consumer theory, the MPI recently has enjoyed widespread use in a production context, in which multiple but cardinally measurable outputs replace scalar-valued but ordinally measurable utility.

Malmquist index has a number of desirable features. It does not require input prices or output prices in their construction, which makes it particularly useful in situations in which prices are distorted or non-existent. It does not require a behavioral assumption such as cost minimization or profit maximization, which makes it useful in situations in which producers' objectives differ, either are unknown or are unachieved. Besides, the MPI is easy to compute, as Färe, Grosskopf, Lindgren and Roos (1992) have demonstrated, and its various decompositions provide insight into the sources of productivity change.

The MPI is based on the concept of production function. Let N inputs $x^t \in R_+^N$ be used to produce M outputs $y^t \in R_+^M$ at time period $t = 1, ..., T$. The set of production possibilities at time t is denoted by

$$\Phi^t = \{(x^t, y^t) \mid x^t \text{ can produce } y^t\}.$$

The upper boundary of Φ^t is sometimes referred to as the *production technology* or the *production frontier*.[1] It is given by the intersection of Φ^t and the closure of its complement.

Following Shephard (1970), standard assumptions on Φ^t are

1) Φ^t is convex;

2) no "free lunches" allowed (i.e. one cannot obtain some output with no input: $\forall y \geq 0 : y \neq 0 \Rightarrow (0, y) \notin \Phi^t$);

3) strong disposability of both inputs and outputs (one can always produce smaller amount of output using the same input; one can always increase the input in order to attain the same output level);

See also Färe (1988) for more details on these assumptions.

A functional representation of production technology is provided by the output distance function

$$D^t(x^s, y^s) \equiv \inf\{\theta > 0 \mid (x^s, y^s/\theta) \in \Phi^t\} \tag{1.1}$$

for the given production unit at time s, relative to the technology existing at time t (see

[1] For simplicity, only output oriented case will be considered here. The same could be done for the input oriented case.

Shephard (1970))[2]. This function gives a normalized measure of distance from the production unit's position in the input-output space at time s to the boundary of the production set at time t in the hyperplane where inputs remain constant. In other words, this function is the reciprocal of proportional increase of the output, given certain amount of input to achieve the efficient frontier. In particular, if $t = s$, then $D^t(x^t, y^t) \leq 1$ and it measures the efficiency relative to the contemporaneous technology. The value $D^t(x^t, y^t) = 1$ indicates that the production unit is technically efficient (since it is on the boundary (frontier) of the production set).

Define now the set V^t as the convex cone (with vertex at the origin) spanned by Φ^t. Analogously to the output distance function with respect to Φ^t, the distance function

$$\Delta^t(x^s, y^s) \equiv \inf\{\theta > 0 | (x^s, y^s/\theta) \in V^t\} \tag{1.2}$$

gives a normalized measure of distance from the given production unit's position in the input/output space at time s to the boundary of V^t in the hyperplane where inputs remain constant. If the technology exhibits constant returns to scale everywhere, then $\Phi^t = V^t$, otherwise $\Phi^t \subset V^t$.

Both output distance functions with respect to Φ^t and with respect to V^t are used in definition and decomposition of the MPI.

Färe, Grosskopf, Norris and Zhang (1994) define the Malmquist productivity index as a geometrical mean of relative productivity changes from time t to time $t+1$ (with respect to the upper boundary of V^t and V^{t+1}):

$$\Pi^{t,t+1} = \left(\frac{\Delta^t(x^{t+1}, y^{t+1})}{\Delta^t(x^t, y^t)} \cdot \frac{\Delta^{t+1}(x^{t+1}, y^{t+1})}{\Delta^{t+1}(x^t, y^t)} \right)^{1/2}. \tag{1.3}$$

It is useful to further decompose that index as:

$$\begin{aligned} \Pi^{t,t+1} &= \frac{\Delta^{t+1}(x^{t+1}, y^{t+1})}{\Delta^t(x^t, y^t)} \cdot \left(\frac{\Delta^t(x^{t+1}, y^{t+1})}{\Delta^{t+1}(x^{t+1}, y^{t+1})} \cdot \frac{\Delta^t(x^t, y^t)}{\Delta^{t+1}(x^t, y^t)} \right)^{1/2} \\ &=: \Delta\text{Eff}^{t,t+1} \cdot \Delta\text{Tech}^{t,t+1}. \end{aligned}$$

In that decomposition, $\Delta\text{Eff}^{t,t+1}$ measures the change in technical effciency over the two periods (i.e. whether or not the unit is getting closer to its efficiency frontier over time), and $\Delta\text{Tech}^{t,t+1}$ measures the change in technology over the two time periods (i.e. whether or not the frontier is shifting out over time). Values of either of these components greater than 1 suggest improvement, while values less than 1 suggest the opposite. Analogously, if the MPI exceeds unity, there has been an improvement in productivity between periods t and $t+1$, while values less than 1 suggest the converse.

Unfortunately, the index defined in the above mentioned decomposition can only measure productivity change if the underlying, true technology exhibits constant returns to

[2]We know that there exist pairs $(x, y) \in R_+^N \times R_+^M$ such that the distance $D^t(x, y)$ given by (1.1) is not defined. Through this paper, when using this notation, we implicitly assume that the quantity is well-defined. Note that the output distance $\Delta^t(x, y)$ given by (1.2) is always well-defined.

scale everywhere (i.e. $\Phi^t = V^t$). Often, this is not the case, since we do not know the true shape of the technology. Therefore, the enhanced decomposition suggested by Simar and Wilson (1998) is called for:

$$\Pi^{t,t+1} = \left(\frac{D^{t+1}(x^{t+1}, y^{t+1})}{D^t(x^t, y^t)}\right) \cdot \left(\frac{\Delta^{t+1}(x^{t+1}, y^{t+1})/D^{t+1}(x^{t+1}, y^{t+1})}{\Delta^t(x^t, y^t)/D^t(x^t, y^t)}\right) \cdot$$

$$\cdot \left(\frac{D^t(x^{t+1}, y^{t+1})}{D^{t+1}(x^{t+1}, y^{t+1})} \cdot \frac{D^t(x^t, y^t)}{D^{t+1}(x^t, y^t)}\right)^{1/2} \cdot$$

$$\cdot \left(\frac{\Delta^t(x^{t+1}, y^{t+1})/D^t(x^{t+1}, y^{t+1})}{\Delta^{t+1}(x^{t+1}, y^{t+1})/D^{t+1}(x^{t+1}, y^{t+1})} \cdot \frac{\Delta^t(x^t, y^t)/D^t(x^t, y^t)}{\Delta^{t+1}(x^t, y^t)/D^{t+1}(x^t, y^t)}\right)^{1/2}$$

$$= \Delta\text{PureEff}^{t,t+1} \cdot \Delta\text{Scale}^{t,t+1} \cdot \Delta\text{PureTech}^{t,t+1} \cdot \Delta\text{ScaleTech}^{t,t+1} \tag{1.4}$$

where $\Delta\text{PureEff}^{t,t+1}$ measures the change in relative efficiency (i.e. the change in how far observed production is from the real maximum potential production) between times t and $t+1$, $\Delta\text{PureTech}^{t,t+1}$ captures the shift in technology between these two time periods, $\Delta\text{Scale}^{t,t+1}$ and $\Delta\text{ScaleTech}^{t,t+1}$ are changes in scale efficiency and scale technology between times t and $t+1$ respectively (see Simar and Wilson (1998) for further details).

In practice, we often observe the productivity units (i.e. input/output set) over time. On the basis of these observations we can estimate the productivity changes for available years, measured in terms of MPI. Often, it is of interest to forecast the MPI for coming year(s). Until now, there are no forecasting methods available in the econometric literature, except for the "naive" ones e.g. using the geometrical mean of the MPI's for some previous years as a prediction of the MPI for the coming year. However, these approaches are purely static, and hide a potentially valuable information given by the evolution of productivity over time. The aim of this paper is to develop a new, dynamical approach for forecasting the MPI, taking into account the behavior of productivity over time.

Several approaches are possible for a dynamic analysis of productivity index. In the next section, we argue that for the sake of forecasting, a desirable characteristic of an index is the circularity property. Because MPI is not circular, we also suggest a new decomposition of the index, in which time-varying indices are circular, making it possible to forecast the MPI by forecasting each "circularized" component separately, and combining the results afterwards.

If one wants to make inference on MPI and construct prediction intervals, bootstrap often appears to be the most attractive possibility. Section 3 presents how smooth bootstrap in that case should be adapted to our situation, including possible time-dependent

structure of the data. Since we have panel data with possible serial correlation (e.g. present performance of the country is influenced by its past performance), a bivariate kernel is used to estimate the joint density of distance functions. One approach to compute confidence intervals is called smooth bootstrap on correlated couples, and was introduced by Simar and Wilson (1999). In Section 4, we further elaborate the smooth bootstrap, allowing for three or more subsequent time periods to be correlated (that leads us to the smooth bootstrap on n-tuples) with no increase of the computational complexity. We compare its performance with the smooth bootstrap on correlated couples by means of MC simulations, using different measures. The results suggest that one indeed benefits from both circularity property and taking the temporal correlation into account.

Section 5 provides an empirical illustration on an economic panel data obtained for 17 OCDE countries. A productivity index is constructed using the GDP, labor and capital stock of the considered countries over the years 1979–1990. This paper provides the first approach to perform a forecast of this important economic index. We compare our prediction with the point estimates of productivity change over the period 1985–1990. Among the forecasters we analyse in this paper, we show that the dynamic approach based on the new decomposition into circular components performs better.

2 Forecasting the MPI

Suppose that we have collected the input-output data for a given production unit up to the year T, and estimated the productivity indices between time t and time $t+1$, denoted $\Pi^{t,t+1}$ for $t = 1, ..., T-1$. Now we would like to forecast the performance of this production unit for the next year $T+1$, that is, to forecast $\Pi^{T,T+1}$.

2.1 Circularity

Starting to think on a consistent way of forecasting a bilateral index such as Malmquist Productivity Index, we face two basic questions: (1) What happens to the index when T, the horizon of observations, tends to infinity, and (2) What happens between two time points of observations. We give an answer to these questions in what follows.

In order to fit a model to the data, and to build a statistically meaningful theory, we would like that: the more data we use, the more accurate forecasts of the MPI should we obtain. It means that if we have an index for the comparison of productivity between time periods t and $t+1$, and between $t+1$ and $t+2$, we must be able to establish a productivity comparison between time periods t and $t+2$ via the arbitrary third time period, $t+1$. As it was shown by Førsund (2002), "circularity" is such a desired property that perfectly suits our needs as a "connector" for indices, and therefore can be used further as a necessary requirement for consistent modelling of a bilateral (i.e. comparing two entities) index.

A bilateral index $I^{t,s}$ is called *circular* if and only if

$$I^{t,t+2} = I^{t,t+1} \cdot I^{t+1,t+2}, \forall t = 1, ..., T-2.$$

Suppose that we have collected data describing the activity of some company for the T time periods. On the basis of these data we have constructed a time sequence of bilateral indices: $\{I_T^{t,t+1} | t = 1, ..., T-1\}$. Here, $I_T^{t,t+1}$ denotes a bilateral index comparing performance of the company during t-th time period to its performance during $(t+1)$-th time period on the basis of data collected for T time periods.

Now, the usual first step when making forecasts based on the time sequence of bilateral indices $\{I_T^{t,t+1} | t = 1, ..., T-1\}$ would be to fit some statistical model to the data. For example, we could use the simplest tentative model:

$$I_T^{t,t+1} = m(t) + e_t$$

where $m(t)$ is a deterministic time-varying function, which is strictly positive, and e_t is a zero-mean stationary process.

In order to construct a forecasting theory, it is necessary to add some regularity assumptions on the function $m(t)$. For example, we can assume the function $m(t)$ to be nearly constant along the intervals of a certain length τ. However, this approach turns out to be very restrictive, since it follows that the function $m(t)$ can be only estimated using exactly T observations. Moreover, when the number of observations T increases, we do not get any substantial improvement of the estimator of $m(t)$ over τ; therefore, we cannot use asymptotic theory for the derivation of usual statistical properties of estimators (such as consistency, efficiency, etc.) and thus, we cannot compare different estimators.

To overcome this drawback, it is useful to write m as a function of the rescaled time t/T , i.e.

$$I_T^{t,t+1} = m(\frac{t}{T}) + e_t, \quad t = 1, ..., T-1$$

where m is now a function over the interval $[0, 1)$. Rescaling is a standard technique in nonparametric regression models (e.g. Fan and Gijbels (1996)). It is also used to get consistent estimates in the context of nonstationary time-series (see Van Bellegem and von Sachs (2004)).

Another natural requirement of an index is the possibility to connect indices based on the data obtained at different frequencies. In other words, if the time sequence of indices is constructed from more frequently observed data (e.g. semestrial data $I^{t,t+1}, t = 1, ..., 2T-1$ observed during $2T$ semesters), the information contained in such time sequence should allow us to reconstruct (with a help of some mapping) the time sequence of indices which are based on less frequently observed data (e.g. yearly data $J^{t',t'+1}, t' =' 1, ..., T-1$ observed during T years).

By rescaling the function m to the interval $[0, 1)$ and using the circularity property, the index can be computed at any time scale. Circularity is a natural interpretable way of connecting indices constructed from data which were observed with different frequencies.

One of the important practical implications of circularity property is the possibility to recover missing values in a time sequence of indices, with a help of a time sequence of indices constructed from the most frequently observed data.

2.2 Decomposition of MPI into circular components

Thus, circularity is a necessary requirement, if we want to forecast the bilateral index, such as MPI. Unfortunately, except for some special cases [3], the Malmquist productivity index is not circular (see Førsund (2002)). In this subsection we are going to introduce a multiplicative decomposition of MPI into circular and stationary factors.

Let us take a closer look at the decomposition (1.4). Circularity of $\Delta\text{PureEff}^{t,t+1}$ and $\Delta\text{Scale}^{t,t+1}$ is easily verified, whereas two other terms $\Delta\text{PureTech}^{t,t+1}$ and $\Delta\text{ScaleTech}^{t,t+1}$ are not circular.

Let us take a closer look at the term

$$\begin{aligned}\Delta\text{PureTech}^{t,t+1} &= \left(\frac{D^t(x^{t+1},y^{t+1})}{D^{t+1}(x^{t+1},y^{t+1})}\cdot\frac{D^t(x^t,y^t)}{D^{t+1}(x^t,y^t)}\right)^{1/2}\\ &=: (\Delta\text{PT}^{t+1}_{t,t+1}\cdot\Delta\text{PT}^{t}_{t,t+1})^{1/2}. \qquad (2.1)\end{aligned}$$

The first factor in the brackets is the relative change of distance to the "true" frontier for a point fixed at time $t+1$, the second term is, respectively, the relative change of distance to the "true" frontier for a point fixed at time t. $\Delta\text{PureTech}^{t,t+1}$ is a geometric mean of these two relative changes. Note that (if the production unit is fixed at times t, or $t+1$ respectively) each one of these terms is circular:

$$\Delta\text{PT}^{t}_{t,t+2} = \Delta\text{PT}^{t}_{t,t+1}\cdot\Delta\text{PT}^{t}_{t+1,t+2},$$

$$\Delta\text{PT}^{t+1}_{t,t+2} = \Delta\text{PT}^{t+1}_{t,t+1}\cdot\Delta\text{PT}^{t+1}_{t+1,t+2}.$$

That inspires the idea of proceeding to the dynamical analysis of Malmquist productivity index by forecasting each of the circular component separately, as follows:

For a given production unit of interest, operating at levels (x^t,y^t) for different values of t, $(t=1,...,T)$ consider a set of components from the decomposition (2.1) : $\Delta\text{PT}^{t}_{s,s+1}$, where $s=1,...,T-1; t=1,...,T$. These can be organized as a table (see Table 1).

[3] MPI is circular (at least) in two special cases: 1)when the production unit is constant over time; 2) when the production technology (frontier) is constant over time.

Shift in years	(x^1, y^1) fixed	(x^2, y^2) fixed	...	(x^T, y^T) fixed	Forecast
$1, 2$	$\Delta PT^1_{1,2}$	$\Delta PT^2_{1,2}$	...	$\Delta PT^T_{1,2}$	$\Delta PT^{T+1}_{1,2}$
$2, 3$	$\Delta PT^1_{2,3}$	$\Delta PT^2_{2,3}$	...	$\Delta PT^T_{2,3}$	$\Delta PT^{T+1}_{2,3}$
$3, 4$	$\Delta PT^1_{3,4}$	$\Delta PT^2_{3,4}$	...	$\Delta PT^T_{3,4}$	$\Delta PT^{T+1}_{3,4}$
...	...	...	...	...	...
$T-1, T$	$\Delta PT^1_{T-1,T}$	$\Delta PT^2_{T-1,T}$	...	$\Delta PT^T_{T-1,T}$	$\Delta PT^{T+1}_{T-1,T}$
Forecast $T, T+1$	$\Delta PT^1_{T,T+1}$	$\Delta PT^2_{T,T+1}$	...	$\Delta PT^T_{T,T+1}$	$\Delta PT^{T+1}_{T,T+1}$

Table 1: The forecasted technological change $\Delta\widehat{\text{PureTech}}^{T,T+1} = (\Delta\widehat{\text{PT}}^{T}_{T,T+1} \cdot \Delta\widehat{\text{PT}}^{T+1}_{T,T+1})^{1/2}$ where $\Delta\widehat{\text{PT}}^{s}_{t,t+1} = \frac{\hat{D}^t(x^s,y^s)}{\hat{D}^{t+1}(x^s,y^s)}$. The "hats" are omitted in the table for convenience of reading.

Here, the sequence

$$\Delta\text{PT}^t_{t,t+1}, \Delta\text{PT}^t_{t+1,t+2}, \Delta\text{PT}^t_{t+2,t+3}, ... \tag{2.2}$$

correspond to the table's columns, and the sequence

$$\Delta\text{PT}^t_{t,t+1}, \Delta\text{PT}^{t+1}_{t,t+1}, \Delta\text{PT}^{t+2}_{t,t+1}, ... \tag{2.3}$$

correspond to the table's rows.

Sequence 2.2, is a sequence of circular indices (as it was shown above), which allows us to forecast the index columnwise. For sequence 2.3, only the production unit changes over time, while the relative change of distance to the "true" frontier remains the same. Therefore it is not a bilateral index, which makes the second sequence a usual time series, and allows forecasting along the table's rows.

Since the term

$$\Delta\text{ScaleTech}^{t,t+1} = \left(\frac{\Delta^t(x^{t+1}, y^{t+1})/D^t(x^{t+1}, y^{t+1})}{\Delta^{t+1}(x^{t+1}, y^{t+1})/D^{t+1}(x^{t+1}, y^{t+1})} \cdot \frac{\Delta^t(x^t, y^t)/D^t(x^t, y^t)}{\Delta^{t+1}(x^t, y^t)/D^{t+1}(x^t, y^t)} \right)^{1/2}$$

has the same structure as the term $\Delta\text{PureTech}^{t,t+1}$, we treat this term in a similar way, which leads us to the two analogous sequences of indices, i.e.:

$$\frac{\Delta^t(x^t, y^t)/D^t(x^t, y^t)}{\Delta^{t+1}(x^t, y^t)/D^{t+1}(x^t, y^t)}, \frac{\Delta^t(x^{t+1}, y^{t+1})/D^t(x^{t+1}, y^{t+1})}{\Delta^{t+1}(x^{t+1}, y^{t+1})/D^{t+1}(x^{t+1}, y^{t+1})}, \frac{\Delta^t(x^{t+2}, y^{t+2})/D^t(x^{t+2}, y^{t+2})}{\Delta^{t+1}(x^{t+2}, y^{t+2})/D^{t+1}(x^{t+2}, y^{t+2})}, ...$$

$$\frac{\Delta^t(x^t, y^t)/D^t(x^t, y^t)}{\Delta^{t+1}(x^t, y^t)/D^{t+1}(x^t, y^t)}, \frac{\Delta^{t+1}(x^t, y^t)/D^t(x^t, y^t)}{\Delta^{t+2}(x^t, y^t)/D^{t+1}(x^t, y^t)}, \frac{\Delta^{t+2}(x^t, y^t)/D^t(x^t, y^t)}{\Delta^{t+3}(x^t, y^t)/D^{t+1}(x^t, y^t)}, ...$$

These two sequences can as well be organized in a table similar to the Table 1, and the forecasting procedure is analogous to the one for $\Delta\text{PureTech}^{t,t+1}$.

2.3 Forecasting the MPI

Since our objective is to forecast the performance of a given production unit in terms of its productivity, we would like to forecast the Malmquist productivity index showing productivity changes from the year T to the year $T+1$, based on the available data for the time span $t = 1, ..., T$. Since the real production frontier and, thus, the MPI's $\Pi^{t,t+1}$ are unknown, we will use the sequence of estimates. The DEA estimation approach used here, is based on Farrell's (1957) ideas, and involves measurement of efficiency for a given production unit relative to the boundary of the convex hull of the input-output set. More precisely, the production frontier at time t is estimated as

$$\hat{\Phi}^t = \{(x,y) \in R^{N+M} \mid y \leq \sum_{i=1}^{n} \gamma_i y_i^t;\ x \geq \sum_{i=1}^{n} \gamma_i x_i^t;\ \sum_{i=1}^{n} \gamma_i = 1; \gamma_j \geq 0, j = 1, ..., n\}.$$

The convex cone spanned by $\hat{\Phi}^t$ is denoted as $\hat{V}^t$:

$$\hat{V}^t = \{(x,y) \in R^{N+M} \mid y \leq \sum_{i=1}^{n} \gamma_i y_i^t;\ x \geq \sum_{i=1}^{n} \gamma_i x_i^t;\ \gamma_j \geq 0, j = 1, ..., n\}.$$

The output distance functions $\hat{D}^t(x^s, y^s)$ and $\hat{\Delta}^t(x^s, y^s)$ (and thus, the MPI) are estimated then with respect to the estimated frontier $\hat{\Phi}^t$ and of $\hat{V}^t$ respectively:

$$\hat{D}^t(x^s, y^s) \equiv \inf\{\theta > 0 \mid (x^s, y^s/\theta) \in \hat{\Phi}^t\}, \tag{2.4}$$

$$\hat{\Delta}^t(x^s, y^s) \equiv \inf\{\theta > 0 | (x^s, y^s/\theta) \in \hat{V}^t\}.$$

The objective now is to forecast the MPI

$$\Pi^{T,T+1} = \Delta\text{PureEff}^{T,T+1} \cdot \Delta\text{Scale}^{T,T+1} \cdot \Delta\text{PureTech}^{T,T+1} \cdot \Delta\text{ScaleTech}^{T,T+1}.$$

The decomposition given above inspires us for the following steps, suggesting to forecast each term of the decomposition separately:

Step 1. Start by forecasting the circular terms: $\Delta\text{PureEff}^{T,T+1}$ and $\Delta\text{Scale}^{T,T+1}$. The forecasting can be performed by using any suitable time-series method, depending on the context (ARMA, exponential smoothing, etc.). We will forecast by exponential smoothing in the empirical example provided in the Section 5, since data for only a short time period is available.

For instance, $\Delta\text{PureEff}^{T,T+1}$ will be forecasted from the sequence of estimates $\widehat{\Delta\text{PureEff}}^{t,t+1}$ for $t = 1, ..., T-1$.

Step 2. Proceed by forecasting the more complicated term, the change of pure technology,

$$\Delta\text{PureTech}^{T,T+1} = \left(\frac{D^T(x^{T+1}, y^{T+1})}{D^{T+1}(x^{T+1}, y^{T+1})} \cdot \frac{D^T(x^T, y^T)}{D^{T+1}(x^T, y^T)}\right)^{1/2} = (\Delta\text{PT}_{T,T+1}^{T+1} \cdot \Delta\text{PT}_{T,T+1}^{T})^{1/2}$$

As it was said before, since both terms ($\Delta\mathrm{PT}^{T+1}_{T,T+1}$ and $\Delta\mathrm{PT}^{T}_{T,T+1}$) in the decomposition are circular (for a fixed production unit), and thus can be forecasted, the main idea here is to treat both terms separately. Indeed, they can be forecasted from the sequence of estimates

$$\Delta\widehat{\mathrm{PT}}^{s}_{t,t+1} = \frac{\hat{D}^t(x^s, y^s)}{\hat{D}^{t+1}(x^s, y^s)}$$

for $s = 1, ..., T; t = 1, ..., T-1$. ($\Delta\widehat{\mathrm{PT}}^{s}_{t,t+1}$ is the estimate of the change in technology between two consecutive years having fixed the year s (i.e., the change in technology is estimated from the perspective of the point (x^s, y^s)).)

Table 1 summarizes the forecasting mechanism. Let us take a closer look at this table. Of interest here are the last two entries in the lower row, since their geometrical mean gives the desired forecast of technological changes for the years $T, T+1$.

First, note that every column in this table has the circularity property, thus allowing us to perform a statistically justified forecasting.

The lower row is obtained by forecasting the corresponding columns (we use exponential smoothing here). In such a way, the forecasting of the column with T-th year fixed gives us the desired term $\Delta\mathrm{PT}^{T}_{T,T+1}$. Afterwards, forecasting the last row, we obtain the other term $\Delta\mathrm{PT}^{T+1}_{T,T+1}$. [4]

Step 3. Similar procedure as before is performed in order to find the estimate $\Delta\mathrm{ScaleTech}^{T,T+1}$, which is decomposed as

$$\Delta\mathrm{ScaleTech}^{T,T+1} =: (\Delta\mathrm{ScT}^{T+1}_{T,T+1} \cdot \Delta\mathrm{ScT}^{T}_{T,T+1})^{1/2}.$$

Each term

$$\Delta\mathrm{ScT}^{T+1}_{T,T+1} = \frac{\Delta^T(x^{T+1}, y^{T+1})/D^T(x^{T+1}, y^{T+1})}{\Delta^{T+1}(x^{T+1}, y^{T+1})/D^{T+1}(x^{T+1}, y^{T+1})}$$

and

$$\Delta\mathrm{ScT}^{T}_{T,T+1} = \frac{\Delta^T(x^{T}, y^{T})/D^T(x^{T}, y^{T})}{\Delta^{T+1}(x^{T}, y^{T})/D^{T+1}(x^{T}, y^{T})}$$

is circular, if we look at the performance of the productivity unit of interest from a fixed-year perspective. Therefore, the estimate $\widehat{\Delta\mathrm{ScaleTech}}^{T,T+1}$ is obtained exactly in the same way as before, using the table analogous to the Table 1.

Step 4. The final forecast of the Malmquist index is given by the product of all the indices forecasted above:

$$\hat{\Pi}^{T,T+1} = \widehat{\Delta\mathrm{PureEff}}^{T,T+1} \cdot \widehat{\Delta\mathrm{Scale}}^{T,T+1} \cdot \widehat{\Delta\mathrm{PureTech}}^{T,T+1} \cdot \widehat{\Delta\mathrm{ScaleTech}}^{T,T+1}.$$

[4] The term $\Delta\mathrm{PT}^{T+1}_{T,T+1}$ can be obtained in other way: first by obtaining the last column (again, using exponential smoothing), then by forecasting this column in order to obtain the entry $\Delta\mathrm{PT}^{T+1}_{T,T+1}$. We will denote these two methods by "Dynamical 1" and "Dynamical 2". Later we will compare the results obtained by both methods and observe that they do not differ significantly.

3 Construction of forecasting intervals

3.1 Bootstrapping Malmquist Index

If one wants to make some inferences on Malmquist productivity indices, bootstrap often appears to be the most attractive possibility. More specifically, after having forecasted the indices of productivity change $\hat{\Pi}^{T,T+1}$ between two consecutive time periods T and $T+1$ for each unit i, we would like to obtain the prediction intervals for the forecasted values. Smoothed bootstrap in that case should be adapted to our situation, including possible time-dependent structure of the data.

Bootstrapping Malmquist index involves replicating this data-generating process, generating a sufficiently large number B of pseudo-samples

$$X_b^* = \{(x_{it}^*, y_{it}^*) | i = 1, ..., N; t = 1, ..., T\};\ b = 1, ..., B$$

and applying the original forecasting procedure to these pseudo-samples, yielding in such a way the bootstrap empirical distributions for Malmquist indices and its components.

If we want to obtain consistent bootstrap estimates of the confidence intervals, we need to replicate the data-generating process in a consistent way as well. However, simple resampling from the empirical distribution of the data (i.e. resampling from the original set of input and outputs, or from the set of distance function estimates) leads to inconsistent bootstrap estimation of the confidence intervals (see Simar and Wilson (2000) for details).

Smooth bootstrap procedure is one of the solutions to overcome this problem and yields consistent estimates. Since we have panel data with the possibility of serial correlation (i.e. present performance of the country is influenced by its past performance), a bivariate kernel (instead of univariate kernel as in the case of a single cross-section of the data) should be used to estimate the joint density of distances $\{(\hat{D}_i^t, \hat{D}_i^{t+1})\}_{i=1}^n$ for each time period t. That estimated bivariate density is used in the bootstrap approach of Simar and Wilson (1999) to simulate the distance function values, a pseudo input/output set, all the components in the decomposition of MPI, and thus, MPI itself for a given production unit. However, because the distances $\hat{D}_i^t$ are potentially correlated over time, we develop in the following an extension of the smooth bootstrap that captures a longer dependence in the data, without any significant additional computational cost.

3.2 Smooth bootstrap on correlated triples

The significant drawback of the previous method (e.g. bootstrapping on correlated pairs) is that the temporal correlation of only two time periods is taken into account. Below we will show that it is possible to extend the method allowing for three subsequent time periods to be correlated. We proceed analogously to the methodology described by Simar and Wilson (1999).

First, consider three (instead of two) $(N \times 1)$ vectors of output distance functions:

$$A = [\hat{D}_1^t, ..., \hat{D}_n^t]'$$

$$B = [\hat{D}_1^{t+1}, ..., \hat{D}_n^{t+1}]'$$

and

$$C = [\hat{D}_1^{t+2}, ..., \hat{D}_n^{t+2}]'.$$

Then, reflect the distance function values in A, B and C about the boundaries (now, there are three boudaries in three-dimensional space $\mathbb{R}^3$), yielding the $(8n \times 3)$ matrix

$$\Delta = \begin{bmatrix} A & B & C \\ 2-A & B & C \\ 2-A & 2-B & C \\ A & 2-B & C \\ A & B & 2-C \\ 2-A & B & 2-C \\ 2-A & 2-B & 2-C \\ A & 2-B & 2-C \end{bmatrix}$$

The corresponding temporal correlation structure is measured by four estimated covariance matrices where

$$\hat{\Sigma}_1 = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} & \hat{\sigma}_{13} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 & \hat{\sigma}_{23} \\ \hat{\sigma}_{13} & \hat{\sigma}_{23} & \hat{\sigma}_3^2 \end{bmatrix}$$

measures the correlation of the original data $[A\ B\ C]$ (and $[2-A\ 2-B\ 2-C]$),

$$\hat{\Sigma}_2 = \begin{bmatrix} \hat{\sigma}_1^2 & -\hat{\sigma}_{12} & -\hat{\sigma}_{13} \\ -\hat{\sigma}_{12} & \hat{\sigma}_2^2 & \hat{\sigma}_{23} \\ -\hat{\sigma}_{13} & \hat{\sigma}_{23} & \hat{\sigma}_3^2 \end{bmatrix}$$

measures the correlation of the original data $[2-A\ B\ C]$ (and $[A\ 2-B\ 2-C]$),

$$\hat{\Sigma}_3 = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} & -\hat{\sigma}_{13} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 & -\hat{\sigma}_{23} \\ -\hat{\sigma}_{13} & -\hat{\sigma}_{23} & \hat{\sigma}_3^2 \end{bmatrix}$$

measures the correlation of the original data $[2-A\ 2-B\ C]$ (and $[A\ B\ 2-C]$), and

$$\hat{\Sigma}_4 = \begin{bmatrix} \hat{\sigma}_1^2 & -\hat{\sigma}_{12} & \hat{\sigma}_{13} \\ -\hat{\sigma}_{12} & \hat{\sigma}_2^2 & -\hat{\sigma}_{23} \\ \hat{\sigma}_{13} & -\hat{\sigma}_{23} & \hat{\sigma}_3^2 \end{bmatrix}$$

measures the correlation of the original data $[A\ 2-B\ C]$ (and $[2-A\ B\ 2-C]$).

Further, we generate the random deviates needed for bootstrap. First, we draw randomly with replacement n rows from Δ to form the $(n\times 3)$ matrix $\Delta^* = [\delta_{ij}], i = 1, ..., n,\ j = 1, 2, 3$. Then we compute the matrix

$$\Gamma = (1+h^2)^{-1/2}(\Delta^* + h\epsilon^* - C \begin{bmatrix} \bar{\delta}_{\cdot 1} & 0 & 0 \\ 0 & \bar{\delta}_{\cdot 2} & 0 \\ 0 & 0 & \bar{\delta}_{\cdot 3} \end{bmatrix}) + C \begin{bmatrix} \bar{\delta}_{\cdot 1} & 0 & 0 \\ 0 & \bar{\delta}_{\cdot 2} & 0 \\ 0 & 0 & \bar{\delta}_{\cdot 3} \end{bmatrix}$$

where $\bar{\delta}_{\cdot j} = \frac{1}{n}\sum_{i=1}^{n} \delta_{ij}$ for $j = 1, 2, 3$, and C is an $(n \times 3)$ matrix of ones, and ϵ^* is a draw from normal density with shape Σ_i, $i = 1, ..., 4$, depending from where the row Δ_i^* was drawn.[5]

Finally, the matrix $\Gamma^* = [\gamma_{ij}]$ of simulated distance function values will be given by

$$\gamma_{ij}^* = \begin{cases} \gamma_{ij}, & if\ \ \gamma_{ij} \geq 1, \\ 2 - \gamma_{ij}, & \text{otherwise.} \end{cases}$$

Remark: An extention to the mentioned above method could be bootstrapping on four or more time periods. However, due to computational complexity, we do not consider this case here, leaving it for the future research. It would be necessary then to develop the criterium for the choice of optimal blocksize.

4 Finite sample properties

We examine the finite sample performance of the method via Monte-Carlo experiments. First, we define our Monte-Carlo scenario. In each Monte-Carlo sample, the set of inputs and outputs is generated according to the Cobb-Douglas model

$$Y_{it} = \alpha_t + \beta_t' X_{it} - \eta_{it}, \quad i = 1, ..., n, \quad t = 1, ..., T$$

where $X_{it} \in R^N$ is a N-dimensional vector of logs of inputs x_{it}, Y_{it} is the log of outputs y_{it}, $\alpha_t \in R^N$ and $\beta_t \in R^N$. The difficulty here is to generate the "inefficiency" term η_{it}, which should be non-negative and autocorrelated. The procedure of the input-output set generation is as follows:

[5] Choice of bandwidth h: Pagan and Ullah (1999) suggest taking $(h = (4/2d+1)^{1/d+4} n^{-1/d+4})$ (i.e. in our case, with $d = 2$, $h = 0.92n^{-1/6}$).

1. Choose the sample size n (number of productivity units), time span T, and the number of inputs N (here, for illustration, N =2 is chosen).

2. Generate the regressors according to the bivariate VAR model

$$X_{it} = RX_{i,t-1} + \nu_{it}, \quad \text{where } \nu_{it} \sim I\mathcal{N}(0, \sigma_X^2 I_2)$$

where $\sigma_X = 1$, I_2 is the 2×2 identity matrix, and the matrix

$$R = \begin{bmatrix} 0.95 & 0.01 \\ 0.01 & 0.95 \end{bmatrix}$$

was chosen such that the process is stationary and X_{it} has reasonable variance. We initialize the simulation with $X_{i1} \sim \mathcal{N}_2(0, \sigma_X^2(I_2 - R^2)^{-1})$, and start iterations for $t \geq 2$.

3. Generate α_t, β_t according to the formulae

$$\alpha_t = \alpha + \gamma_1 \frac{t}{T}, \quad \beta_t = \beta + \gamma_2 \frac{t}{T}.$$

In the simulations, we choose $\alpha = 1, \beta = 1, \gamma_1 = 0.01, \gamma_2 = 0.01$.

4. The error term η_{it} of the model is generated as

$$\eta_{it} = \lambda_i + \epsilon_{it},$$

where $\lambda_i \sim Exp(1)$, so that the λ_i are positive. For the ϵ_{it} AR(1) model is adapted, with $\epsilon_{it} = 0.5\epsilon_{i,t-1} + e_{it}$, where $e_{it} \sim Unif[-\lambda_i/2, \lambda_i/2]$, in order to ensure that $\eta_{it} > 0$.

5. Finally, obtain x_{it} and y_{it} by taking exponentials to come back to original coordinates.

Due to computing limitations, the results only for $M = 500$ Monte Carlo replications were obtained. Remember, that our objective is to forecast the MPI between time T and $T-1$, $\Pi_i^{T,T+1}$. Note that we know the true values of $\Pi_i^{t,t+1}$, $i = 1, ..., T-1$, since we know the true production function (we have just simulated it as $Y_{it} = \alpha_t + \beta_t' X_{it} - \eta_{it}$), and hence are able to calculate all the distances to the production frontier.

Then for each sample, based on the simulated input-output set, we can estimate a Malmquist productivity index $\hat{\Pi}_i^{T-1,T}$ and forecast $\hat{\Pi}_i^{T,T+1}$ using the dynamic method. For the forecasted MPI, two confidence intervals were obtained using: a) bootstrap on correlated pairs, b) bootstrap on correlated triples. After repeating this procedure M times, the performance of confidence intervals is evaluated (see Table 2) with a help of averaging their lengths, and estimation of coverage probabilities (proportion of times, over M replications, that the confidence interval contains the true index $\Pi_i^{T-1,T}$).

As it can be seen from the Table 2, the bootstrap on correlated triples performs considerably better in terms of coverage than the bootstrap on correlated pairs. The length of the interval is also reduced, it is especially well seen for the times $T = 30$ (0.6 compared

to 0.45) and for $T = 45$ (0.64 compared to 0.48). This indicates that, using bootstrap on correlated triples, we obtain more precise results that using bootstrap on correlated pairs. This, probably, is due to the fact that the bootstrap on correlated triples takes into account the correlation between three consequent time periods instead of two time periods, as in the case of the bootstrap on correlated pairs, allowing to achieve greater level of preciseness.

	Corr. Pairs		**Corr. Triples**	
T	**Length**	**Coverage**	**Length**	**Coverage**
10	0.76	0.83	0.68	0.88
15	0.54	0.91	0.61	0.94
30	0.6	0.88	0.45	0.93
45	0.64	0.90	0.48	0.93

Table 2: Smooth bootstrap on correlated pairs vs. triples. Performance of confidence intervals: average length of interval and estimation of coverage probabilities, $n = 30$, $T = 10, 15, 30, 45$. Number of Monte-Carlo experiments $M = 500$, bootstrap $B = 1000$.

5 Application on OECD data

In order to illustrate the suggested methodology for forecasting the Malmquist productivity index in details, let us consider the application of this method on a real dataset. The sample of 17 OECD (Organisation for Economic Co-operation and Development) countries for years 1979-1990 was taken from the Penn World Tables (version 5.6) described by Summers and Heston (1991). The countries considered here are: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Norway, Spain, Sweden, the United Kingdom, and the United States. The data was collected as a result of the benchmark studies performed by the International Comparison Program of the United Nations and national-account data.

Our objective is to forecast the performance of a given country in terms of its productivity. The forecasting of productivity indices $\hat{\Pi}^{90,91}$ has been performed for 17 countries using the OCDE dataset.

Variables chosen for comparison are: GDP, labor and capital stock. GDP and capital stock are measured in 1985 international prices. Labor is retrieved from real GDP per worker, and capital is retrieved from capital stock per worker.

In this example, labor and capital stand for input, and GDP - for output. This choice of the variables is quite natural, since a country is usually interested in maximizing its productivity (in terms of GDP), which can be traditionally expressed through the labor-capital correspondence which relates value-added to primary (capital and labour) inputs.

Here we use output-oriented measures, so a value of the Malmquist productivity index greater than one indicate a progress, etc.

In the following we illustrate our method on Japan. Results for other countries can be obtained in a similar way. In other words, we would like to forecast the Malmquist productivity index showing productivity changes for years 1990-1991, based on the available data. We are able to find the estimates for Malmquist productivity indices up to the year 1990:

$$\hat{\Pi}_{Jap}^{79,80}, \hat{\Pi}_{Jap}^{80,81}, ..., \hat{\Pi}_{Jap}^{89,90}$$

where $\hat{\Pi}_{Jap}^{t,t+1}$ denotes how did the productivity of Japan change from the year t to the year $(t+1)$. In order to see the quality of forecasting, let us take a look at Figure 1.

Figure 1: Comparison of 4 methods of forecasting the Malmquist productivity index (Japan). Solid line indicates estimated MPI, dotted line indicates geometrical mean of the MPI for preceding years, upper dashed line denotes the MPI forecasted by exponential smoothing, and lower dashed and dash-dotted lines stand for the MPI forecasted by Dynamical methods 1 and 2, respectively

Here, solid line shows estimated Malmquist indices (for Japan), dotted line indicates Malmquist indices obtained by the "naive" method (geometrical mean of MPI's for preceding years), and lower dashed and dash-dotted lines (which are practically indistinguishable) indicate Malmquist indices forecasted by Dynamical method 1 and Dynamical method 2 respectively.

Upper dashed line indicates Malmquist indices obtained by the exponential smoothing

of the MPIs estimated for preceding years. We can see that starting from years 85-86 (i.e. having more data available for forecasting), the quality of forecasting using dynamical methods is quite satisfactory and much better than the quality of forecasting using the exponential smoothing.

The results corresponding to the Figure 1 (again, for Japan) are listed below at Table 5.

Years	Estimated	Forecasted *"Exp. smoothing"*	Forecasted *"Geom.Mean"*	*Dynamic 1*	*Dynamic 2*
79-80	1.0158				
80-81	1.0182				
81-82	1.0210				
82-83	1.0179				
83-84	0.9981				
84-85	0.9911				
85-86	1.0187	1.0103	1.1022	1.0065	1.0056
86-87	1.0000	1.0115	1.0342	1.0037	1.0014
87-88	0.9726	1.0100	1.0008	0.9896	0.9868
88-89	0.9755	1.0058	1.0472	0.9705	0.9668
89-90	0.9626	1.0027	0.9865	0.9395	0.9369
Median abs.err.		0.0303	0.0342	0.0121	0.0131
Median SE		0.0009	0.0012	0.0001	0.0002
Ind. function		0.4	0.6	1	1

Table 3: Forecasted vs. estimated productivity index for Japan.

In order to compare the 3 methods discussed above, several ways to measure the quality of forecast are suggested. These are

- Median absolute error (MAE): the median value of the set $|\Pi_{Jap}^{85,86} - \hat{\Pi}_{Jap}^{85,86}|, ..., |\Pi_{Jap}^{89,90} - \hat{\Pi}_{Jap}^{89,90}|$, where $\Pi_{Jap}^{j,j+1}$ is the estimated value of the productivity index, and $\hat{\Pi}_{Jap}^{j,j+1}$ is the value of productivity index, forecasted by one of three mentioned methods: static, Dynamic 1, or Dynamic 2.

- Median squared error (MSE): the median value of the set $|\Pi_{Jap}^{85,86} - \hat{\Pi}_{Jap}^{85,86}|^2, ..., |\Pi_{Jap}^{89,90} - \hat{\Pi}_{Jap}^{89,90}|^2$ where $\Pi_{Jap}^{j,j+1}$ is the estimated value of the productivity index, and $\hat{\Pi}_{Jap}^{j,j+1}$ is the value of productivity index, forecasted by one of three mentioned methods: static, Dynamic 1, or Dynamic 2.

- Indicator function $F(t)$, counting the proportion of times the direction of productivity change (i.e less or greater than unity) is predicted correctly

$$F(t) = \frac{\#\{sign(\Pi_i^{j,j+1}) = sign(\hat{\Pi}_i^{j,j+1})\}}{T}$$

whith value of indicator function being the estimated probability to predict the right direction of productivity change.

One can notice that all the measures suggest that, as expected, the dynamic methods perform better, the indicator function shows as well that one might prefer using dynamical approach, since it might be more important for practitioner to predict correctly the direction of future productivity changes,than its numerical value.

In general (over all the countries), the dynamic method predicts correctly the direction of changes in productivity in 86% (Dynamic method 1) and 85% (Dynamic method 2), compared to only 57% of correct predictions on case of using the static method.

The results of forecasting the Malmquist productivity index for the years 89-90 are shown at the Table 4. Here, the second and third columns contain the $\hat{\Pi}^{89,90}$ forecasted by dynamic and static methods respectively (for simplicity, the results only for "dynamic method 1" are given below, since both methods differ insignificantly), and the last column is the estimated productivity index on the base of available data. Here, single asterisks denote significant differences from unity at 95% significance level, determined by whether the estimated bootstrap (on correlated couples) confidence intervals contains unity, and "+"' denotes significant differences from unity at 95% significance level, determined by whether the estimated bootstrap (on correlated triples) confidence intervals contains unity.

Note that the MPI for Ireland is not available, since $\hat{D}^t(x^s, y^s)$, as computed in (2.4), is not defined for that country. Therefore, the estimates for pure efficiency and other terms in the decomposition (1.4) are not available (see also Simar and Wilson (1998)).

Country	Dynamic	Geom. Mean	"True"
Australia	1.0059^{*+}	0.9848	1.0316
Austria	0.9489^{*+}	1.0077	0.9826
Belgium	0.9454	0.9864	0.9696
Canada	0.9945^{*+}	0.9818	0.9987
Denmark	0.9632^{*+}	0.9966	0.9980
Finland	0.9890	0.9711	0.9992
France	0.9603^{*+}	0.9888	0.9868
Germany	0.9418^{*+}	0.9955	0.9659
Greece	1.0025^{*+}	1.0015	1.0031
Italy	0.9734^{*+}	1.0000	0.9853
Japan	0.9369	1.0027	0.9626
Norway	0.9639	0.9845	0.9886
Spain	0.9758^{*+}	1.0027	0.9899
Sweden	0.9746^{+}	0.9861	0.9993
UK	0.9980	0.9974	1.0468
USA	0.9794	0.9963	1.014

Table 4: Forecasted by 3 different methods Malmquist productivity index $\hat{\Pi}^{89,90}$ for 17 OCDE countries.

Ten of the forecasts (about 60 percent) are significantly different from unity. Of these, 8 indicate decreasing productivity; only Australia and Greece show an increase in productivity, and these are small.

	Geom.Mean			Dynamic 1			Dynamic 2		
	MAE	MSE	$F(t)$	MAE	MSE	$F(t)$	MAE	MSE	$F(t)$
Australia	0.0212	0.0004	3	0.0077	0.0001	4	0.0074	0.0001	4
Austria	0.0251	0.0006	2	0.0088	0.0001	3	0.0112	0.0001	3
Belgium	0.0168	0.0003	5	0.0078	0.0001	5	0.0073	0.0001	5
Canada	0.0218	0.0005	4	0.0081	0.0001	4	0.0071	0.0001	4
Denmark	0.0140	0.0002	2	0.0119	0.0001	4	0.0131	0.0002	4
Finland	0.0234	0.0005	4	0.0073	0.0001	4	0.0073	0.0001	4
France	0.0145	0.0002	5	0.0080	0.0001	5	0.0072	0.0001	5
Germany	0.0270	0.0007	3	0.0072	0.0001	5	0.0072	0.0001	5
Greece	0.0067	0.0000	3	0.0120	0.0001	4	0.0109	0.0001	4
Italy	0.0180	0.0003	1	0.0114	0.0001	4	0.0118	0.0001	4
Japan	0.0303	0.0009	2	0.0121	0.0001	5	0.0131	0.0002	5
Norway	0.0239	0.0006	3	0.0072	0.0001	5	0.0072	0.0001	5
Spain	0.0324	0.0010	1	0.0110	0.0001	4	0.0100	0.0001	4
Sweden	0.0032	0.0000	5	0.0105	0.0001	5	0.0072	0.0001	5
UK	0.0232	0.0005	1	0.0037	0.0000	4	0.0036	0.0000	3
USA	0.0107	0.0001	2	0.0092	0.0001	4	0.0118	0.0001	4

Table 5: Quality of forecasting by 3 methods for OECD countries

6 Conclusions

As far as we know, very few relevant (to the dynamic approach for forecasting the MPI) papers exists in the literature on productivity analysis. In this paper, we suggest and investigate a new working procedure of dynamic forecasting of the Malmquist productivity index. The new decomposition of the index into circular and stationary components allows us to forecast it more efficiently than the non-circularized one.

Illustrated by the empirical results obtained from OCDE dataset, the proposed forecasting method works better with respect to several forecasting error criteria, and the difference with the static methods is quite significant. However, the lack of observations (only 11 years available) does not allow us to use more sophisticated forecasting methods, and one could hope for better results when having data for more years.

For the inferences on MPI, the existing method of smooth bootstrap on correlated pairs is extended, taking three or more subsequent time periods into account. As confirmed by the Monte-Carlo simulations, this results in a better performance of the confidence intervals for the forecasted values, both in terms in length and coverage.

References

Caves, D. W., Christensen, L. R. and Diewert, W. E. (1982). Multilateral comparisons of output, input and productivity using superlative index numbers. *Economic Journal*, *92(365)*, 73-86.

Fan, J. and Gijbels, I. (1996). *Local polynomial modelling and its applications*. London: Chapman and Hall.

Färe, R. (1988). *Fundamentals of production theory*. Heidelberg: Springer Verlag.

Färe, R., Grosskopf, S., Lindgren, B. and Roos, P. (1992). Productivity changes in swedish pharmacies 1980-1989: A nonparametric malmquist approach. *Journal of Productivity Analysis*, *3*, 85-101.

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

Førsund, F. R. (2002). On the circularity of the Malmquist productivity index. *ICER working papers*.

Malmquist, S. (1953). Index numbers and indifference curves. *Trabajos de Estadistica*, *4(1)*, 209-242.

Pagan, A. and Ullah, A. (1999). *Nonparametric econometrics*. Cambridge University Press.

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

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

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

Simar, L. and Wilson, P. (2000). Statistical inference in nonparametric frontier models:the state of the art. *Journal of Productivity Analysis*, *13*, 49-78.

Summers, R. and Heston, A. (1991). The Penn World Table (Mark 5): An expanded set of international comparisons, $1950 - 1987$. *Quaterly Journal of Economics*, *106*, 1-41.

Van Bellegem, S. and von Sachs, R. (2004). Forecasting economic time series with unconditional time-varying variance. *International Journal of Forecasting*, *20(4)*, 611-628.