

# A toolbox for calculating and decomposing Total Factor Productivity indices

Bert M. Balk<sup>a</sup>, Javier Barbero<sup>b,c,\*</sup>, José L. Zofío<sup>d,e</sup>

<sup>a</sup>Rotterdam School of Management, Erasmus University, Netherlands

<sup>b</sup>European Commission, Joint Research Centre (JRC), Spain

<sup>c</sup>Oviedo Efficiency Group, University of Oviedo, Spain

<sup>d</sup>Universidad Autónoma de Madrid, Spain

<sup>e</sup>Erasmus Research Institute of Management, Erasmus University Netherlands

## ARTICLE INFO

### Article history:

Received 7 February 2019

Revised 16 September 2019

Accepted 9 November 2019

Available online 27 November 2019

### Keywords:

Total factor productivity

Malmquist

Moorsteen–Bjurek

Fisher

Törnqvist

Data envelopment analysis

## ABSTRACT

**Total Factor Productivity Toolbox** is a new set of functions to calculate the main Total Factor Productivity (TFP) indices and their decompositions, based on Shephard's distance functions, and using Data Envelopment Analysis (DEA) programming techniques. The package includes code for the standard Malmquist, Moorsteen–Bjurek, price-weighted and share-weighted TFP indices, allowing for the choice of orientation (input or output), reference period (base, comparison, geometric mean), returns to scale (variable or constant), and specific decompositions (aggregate, or identifying scale effects, as well as input and output mix effects). Classic definitions of TFP corresponding to the Laspeyres, Paasche, Fisher, or Törnqvist formulas can also be calculated as particular cases. This paper describes the methodology and implementation of the productivity functions in MATLAB. We compare the results corresponding to the different definitions by studying productivity trends in the US agriculture at the individual state level.

© 2019 The Author(s). Published by Elsevier Ltd.

This is an open access article under the CC BY license. (<http://creativecommons.org/licenses/by/4.0/>)

## 1. Introduction

Total factor productivity (TFP) change is an important concept in economics because it measures the ability of firms, industries, and national economies to increase the aggregate volume of outputs they yield, relative to the aggregate volume of inputs they use. TFP constitutes a standard instrument for monitoring and benchmarking observations, and represents the cornerstone in multilateral studies of technological and economic performance, OECD (2001). Defined as the ratio of an output quantity index to an input quantity index, TFP change can be calculated and decomposed in various ways.

TFP measurement has grown in importance over the past decades due to the increasing availability of data to study the productive performance of units, regardless of their market, governmental, or not-for-profit orientation. However, there are many ways to measure TFP depending on whether quantities and prices are available. TFP measurement requires the aggregation of quantities through suitable functions. If prices are available, it is pos-

sible to aggregate output and input quantities by index formulas of Laspeyres, Paasche, Fisher, or Törnqvist. If only quantities are available, the concept of distance function as introduced by Shephard emerges as the building block in the definition of Malmquist and Moorsteen–Bjurek indices. Aggregation then relies on optimization—mathematical programming—techniques such as the non-parametric Data Envelopment Analysis (DEA, Cooper et al., 2007).

There are various ways to decompose TFP change to identify the components. This applies both to the classical definitions using prices as aggregators, and those relying only on quantities through distance functions. These components are (technical) efficiency change, technological change, scale effect, and changes in the mix of inputs and outputs. The various definitions of these terms that have been proposed over the years gave rise to a healthy debate between authors, including (Balk, 2001; 2003; Färe et al., 2008; 1994; Lovell, 2003; Ray and Desli, 1997; Zofío, 2007), and, more recently, (O'Donnell, 2018). Balk and Zofío (2018) examine how to identify those terms that allow a meaningful interpretation and decomposition of TFP in a general framework. Although the present paper is self-contained, the toolbox follows these authors, where more detailed theoretical and empirical considerations can be found. The toolbox implements several functions

\* Corresponding author.

E-mail address: [Javier.BARBERO-JIMENEZ@ec.europa.eu](mailto:Javier.BARBERO-JIMENEZ@ec.europa.eu) (J. Barbero).

URL: <http://www.javierbarbero.net> (J. Barbero)

to calculate the main quantity-only and price-based TFP indices proposed in the literature, along with their associated decompositions. The toolbox allows for a choice of orientation (input or output), reference period (base, comparison, geometric mean), returns to scale (variable or constant), and specific decompositions (aggregate or identifying scale effects, as well as input and output mix effects). The code calculates Shephard's input and output distance functions approximating a production technology empirically through DEA. The toolbox also calculates classic TFP measures that do not rely on distance functions for their definition, including Laspeyres, Paasche, Fisher, and Törnqvist.

Quantity-only TFP indices based on distance functions, corresponding mainly to the Malmquist index, and coupled occasionally with the Moorsteen-Bjurek index, can be found in standard software packages like Stata (StataCorp, 2015); through user-written commands as Lee et al. (2011); in LIMDEP (Econometric Software 2014); in dedicated non-commercial software accompanying academic handbooks such as in Cooper et al. (2007) and Bogetoft and Otto (2011) (implemented in R); in commercial software, including trial versions (O'Donnell, 2011 and Emrouznejad and Thanassoulis, 2011); in free-ware programs (Coelli, 1998); and even in tutorials for spreadsheets (Zhu, 2014). For the classic TFP indices using prices as aggregators, Shehata and Mickael (2015) developed a Stata module to calculate value index numbers, while Coelli (1999) offered a stand-alone program that calculates Fisher and Törnqvist indices, including transitive versions using the EKS method. O'Donnell (2011) expanded this last option to Lowe-type indices under a commercial license. Recently, Dakpo et al. (2018) have coded an R-based productivity package, based on O'Donnell (2011, 2012, 2018). However, while it covers many of the definitions considered in this toolbox, the factors corresponding to technological change, as well as the scale and the input- and output-mix effects vary as result of differences in the underlying theoretical models. As remarked by O'Donnell (2012), there is a potentially infinite number of exhaustive decompositions of a TFP index. For example, one family of these decompositions, identified by Dakpo et al. (2018) and O'Donnell (2011, 2012, 2018), defines technological change globally as the change in maximum productivity between two time periods, while in this toolbox a local measure of technological change is considered. Consequently, the rest of factors must accommodate the numerical difference between the two. However, all these decompositions can be rightly interpreted, while complementing each other. Hence, researchers can rely on one or another depending on their preferred definitions, and compare them using all the available packages. Additionally, as a novelty of this toolbox, we offer the possibility of decomposing share-weighted geometric indices, which is unavailable in previous packages. Finally, a comprehensive review of the available general purpose and dedicated software options for efficiency and productivity analysis can be found in Daraio et al. (2019).

Although these packages implement the main TFP indices, there is a lack of a full set of functions for MATLAB, and none of them includes a complete decomposition of productivity change according to its multiple components. Thus, besides implementing the basic TFP definitions based on quantities and prices in the MATLAB environment, our toolbox calculates a large array of index numbers capturing (technical) efficiency change, technological change, as well as scale effects, and input and output mix effects. Quantities-only Malmquist and Moorsteen-Bjurek indices, as well as price-based Fisher and Törnqvist indices, are decomposed into mutually exclusive factors with meaningful interpretations in terms of economic index number theory. The **Total Factor Productivity Toolbox** introduces a set of functions, covering a wide range of TFP indices, and reports numerical results using a common example to ease comparability and to illustrate their use. The toolbox is available as free software, under the GNU Gen-

eral Public License version 3, and can be downloaded from <http://www.tfptoolbox.com>, with all the supplementary material (data, examples, and source code) to replicate all the results presented in this paper. The toolbox is compatible with the 2017b (8.5) version of MATLAB, (The MathWorks, Inc., 2017). The U.S. agricultural data employed in this article, as well the code used to calculate the various productivity indices are available as supplementary on-line material. The toolbox is also hosted on an open source repository on GitHub.<sup>1</sup>

This paper is organized as follows. Section 2 presents the data structures characterizing the production possibility sets, the structure of the functions, results, and the U.S. agricultural data that is used as real-case study and to illustrate the toolbox. Section 3 covers the Malmquist productivity indices, relying on radial output or input distance functions. We show how these indices can be decomposed into factors with meaningful interpretations, such as technical efficiency change, technological change, scale effect, and input and output mix effect. We also relate and interpret these factors in terms of the output code that is obtained when running the specific functions. Malmquist productivity indices take into consideration only one of the two sides of the production process, output or input. In Section 4 we consider the class of Moorsteen-Bjurek indices, defined as ratio of an output quantity index to an input quantity index, which in turn can be expressed in terms of output and input distance functions, respectively. We present the decomposition of these indices using various alternatives to identify the above effects. If input and output prices are available, it is possible to calculate and decompose classical indices such as Fisher and Törnqvist. The price-weighted productivity indices are considered in Section 5. Rather than multiplying input and output quantities by their prices, one can aggregate individual quantity ratios by a geometric mean, weighing with cost or revenue shares. Section 6 deals with the class of share-weighted indices, whose best known representative is the Törnqvist productivity index. Advanced options, including displaying and exporting results are discussed in Section 2.2. Section 8 concludes.

## 2. Data structures, output tables and empirical analysis

### 2.1. Data structures

Total Factor Productivity change measures the temporal variation in the productive performance of a DMU (decision making unit such as a plant, firm, industry, or economy) over time. In time period  $t = 0, 1, \dots, T$  each DMU transforms a vector of inputs  $\mathbf{x}^{kt} \in \mathbb{R}_{++}^N$  into a vector of outputs  $\mathbf{y}^{kt} \in \mathbb{R}_{++}^M$  ( $k = 1, \dots, K$ ). Piecewise linear approximations of the period-specific technologies can be obtained through Data Envelopment Analysis techniques. Specifically, the constant returns to scale (CRS) production possibility set introduced by Charnes et al. (1978) (CCR), corresponds to  $S^t = \{(\mathbf{x}^t, \mathbf{y}^t) \mid \mathbf{x}^t \geq X^t \lambda, \mathbf{y}^t \leq Y^t \lambda, \lambda \geq \mathbf{0}\}$ , where  $X^t = (\mathbf{x}^{kt})_{k \in \{1, \dots, K\}} \in \mathbb{R}^{N \times K}$  and  $Y^t = (\mathbf{y}^{kt})_{k \in \{1, \dots, K\}} \in \mathbb{R}^{M \times K}$  are matrices, and  $\lambda = (\lambda_1, \dots, \lambda_K)'$  is a semi-positive vector. Alternatively, the variable returns to scale (VRS) production possibility set introduced by Banker et al. (1984) (BCC), corresponds to  $S^t = \{(\mathbf{x}^t, \mathbf{y}^t) \mid \mathbf{x}^t \geq X^t \lambda, \mathbf{y}^t \leq Y^t \lambda, \mathbf{e} \lambda = \mathbf{1}, \lambda \geq \mathbf{0}\}$ , where  $\mathbf{e}$  is a row vector with all elements equal to 1. The only difference with the CCR model is the adjunction of the condition  $\sum_{k=1}^K \lambda_k = 1$ . If input and output prices are available,  $\mathbf{w}^{kt} \in \mathbb{R}_{++}^N$  and  $\mathbf{p}^{kt} \in \mathbb{R}_{++}^M$ , we have the following panel data structure:  $(\mathbf{w}^{kt}, \mathbf{x}^{kt}, \mathbf{p}^{kt}, \mathbf{y}^{kt})$  ( $k = 1, \dots, K; t = 0, 1, \dots, T$ ). Data are managed as regular MATLAB vectors and matrices, constituting the inputs of the estimation functions that are described in what follows.

<sup>1</sup> The address of the repository is <https://github.com/joselzofio/TFPMATLAB>.

All estimation functions return a structure `tfpout` that contains fields with the estimation results as well as the input of the estimation function. Fields can be accessed directly using the dot notation, and the whole structure can be used as an input to other functions that print or export results (e.g., `tfpdisp`). Some of the fields of the `tfpout` structure are the following:<sup>2</sup>

- `X` and `Y`: Contain the input and output quantities, respectively.
- `W` and `P`: Contain the input and output prices, respectively.
- `n`: Number of observations.
- `m` and `s`: Number of input and output variables, respectively.
- `orient`, `period`, `decomp`: Strings containing the orientation of the TFP model, reference period, and decomposition depending on scale and mix effects assumptions.
- `tfp.M`, `tfp.MB`, `tfp.PROD`, and `tfp.GPROD`: Computed Malmquist, Moorsteen–Bjurek, price-weighted, and share-weighted productivity indices.
- `tfp.EC`, and `tfp.TC`: Computed technical efficiency change and technological change factors.
- `tfp.SEC`, `tfp.OME`, `tfp.IME` and `tfp.RTS`: Computed scale effect, output mix effect, input mix effect, and returns to scale factors.
- `names`: Names of the DMUs.

## 2.2. Displaying and exporting results

### 2.2.1. Custom display

Throughout the text, we illustrate how to display output tables by calling the `tfpdisp(out, dispstr)` function after computing a certain decomposition. In the table header appropriate information concerning the estimated model is displayed on the screen. This setting can be changed to display bespoke information by specifying in the `tfpdisp` function the string `dispstr` (display string) as a second parameter.

For example, the default `dispstr` after using the `deatfpgprod` function with the geometric mean option is `names/ tfp.GProd/ tfp.EC/ tfp.TC/ tfp.IME/ tfp.SEC`. The fields displayed in the output table must be separated by a / and include the names corresponding to the field names of the `tfpout` structure. The available fields are those presented in Table 1.

### 2.2.2. Exporting results

Results can easily be exported to various file formats for posterior analysis and sharing. First, the `tfpout` structure should be converted to a MATLAB table data type by using the `tfp2table(out, dispstr)` function with the desired `dispstr`. If the `dispstr` parameter is omitted, the default is used.

```
> tfpgprod_exp = deatfpgprod(X, Y, W, P, 'orient', 'io', ...
'period', 'geomean', 'decomp', 'complete');
> T = tfp2table(tfpgprod_exp)
```

The table can then be exported by using the function `writetable`.<sup>3</sup>

```
> writetable(T, 'tfpgprod_results.csv')
```

## 2.3. Agricultural productivity in the US

To illustrate the toolbox we compare the various productivity trends in US agriculture obtained with the existing TFP def-

**Table 1**

Fields of the `tfpout` structure available for the `dispstr` string.

Common fields	
Field	Data
<code>names</code>	DMU names
<code>EC</code>	Technical efficiency change
<code>TC</code>	Technological change
<b>Productivity indices</b>	
<code>tfp.M</code>	Malmquist productivity index (after <code>deatfpm</code> )
<code>tfp.MB</code>	Moorsteen–Bjurek productivity index (after <code>deatfpm</code> )
<code>tfp.Prod</code>	Price-weighted productivity index (after <code>deatfprod</code> )
<code>tfp.GProd</code>	Share-weighted productivity index (after <code>deatfpgprod</code> )
<b>Output orientation</b>	
<code>tfp.SEC</code>	Scale effect (geometric mean)
<code>tfp.OME</code>	Output mix effect (geometric mean)
<code>tfp.SEC_100</code>	$SEC_0^c(x^1, x^0, y^0)$
<code>tfp.OME_110</code>	$OME^c(x^1, y^1, y^0)$
<code>tfp.SEC_101</code>	$SEC_0^c(x^1, x^0, y^1)$
<code>tfp.OME_010</code>	$OME^c(x^0, y^1, y^0)$
<b>Input orientation</b>	
<code>tfp.SEC</code>	Scale effect (geometric mean)
<code>tfp.IME</code>	Input mix effect (geometric mean)
<code>tfp.IME_100</code>	$IME^c(x^1, x^0, y^0)$
<code>tfp.SEC_110</code>	$SEC_1^c(x^1, y^1, y^0)$
<code>tfp.IME_101</code>	$IME^c(x^1, y^0, y^1)$
<code>tfp.SEC_010</code>	$SEC_1^c(x^0, x^1, y^0)$

initions. The data has been collected and tabulated by the Economic Research Section (ERS) of the United States Department of Agriculture (USDA) in an effort to study long term productivity trends. It corresponds to a subset of the state-level tables including price indices and implicit quantities of farm outputs and inputs. It is readily available in MATLAB format as supplemental on-line material to this article. We calculate productivity growth between 1960 and 2004, corresponding to the first and last available years in the dataset. The full data, corresponding to Table #23, can be downloaded from <https://www.ers.usda.gov/data-products/agricultural-productivity-in-the-us/>. Our dataset consists of three outputs (livestock, crops, and other farm related output), and four inputs (capital, land, labor, and intermediate inputs). More information on the data, including descriptive statistics and a discussion of the agricultural productivity growth in the US can be found in the above link. Due to its comprehensiveness and reliability, this dataset has been used over the years by many authors, whose results are complemented with those obtained by solving the models included in this toolbox; see, to cite but a few, Färe et al. (2008), Ball et al. (2001), and Zofio and Lovell (2001). Once downloaded, data on input and output quantities and prices can be brought into the workspace by running the following code included in the example file:

```
> load 'DataAgriculture'
> X1960 = [CAPITAL1960, LAND1960, LABOR1960, INTINP1960];
> X2004 = [CAPITAL2004, LAND2004, LABOR2004, INTINP2004];
> Y1960 = [LS1960, CROPS1960, FARMOUT1960];
> Y2004 = [LS2004, CROPS2004, FARMOUT2004];
> W1960 = [WCAPITAL1960, WLAND1960, WLABOR1960, WINTINP1960];
> W2004 = [WCAPITAL2004, WLAND2004, WLABOR2004, WINTINP2004];
> P1960 = [PLS1960, PCROPS1960, PFARMOUT1960];
> P2004 = [PLS2004, PCROPS2004, PFARMOUT2004];
> X = X./ 1000; Y = Y./ 1000;
```

## 3. Malmquist productivity index

When only quantities are available the most popular measurement instrument is the Malmquist productivity index (MPI). The

<sup>2</sup> For a full list see the help of the function typing `help tfpout` in MATLAB.

<sup>3</sup> See the official documentation for this function at <http://www.mathworks.com/help/matlab/ref/writetable.html>.

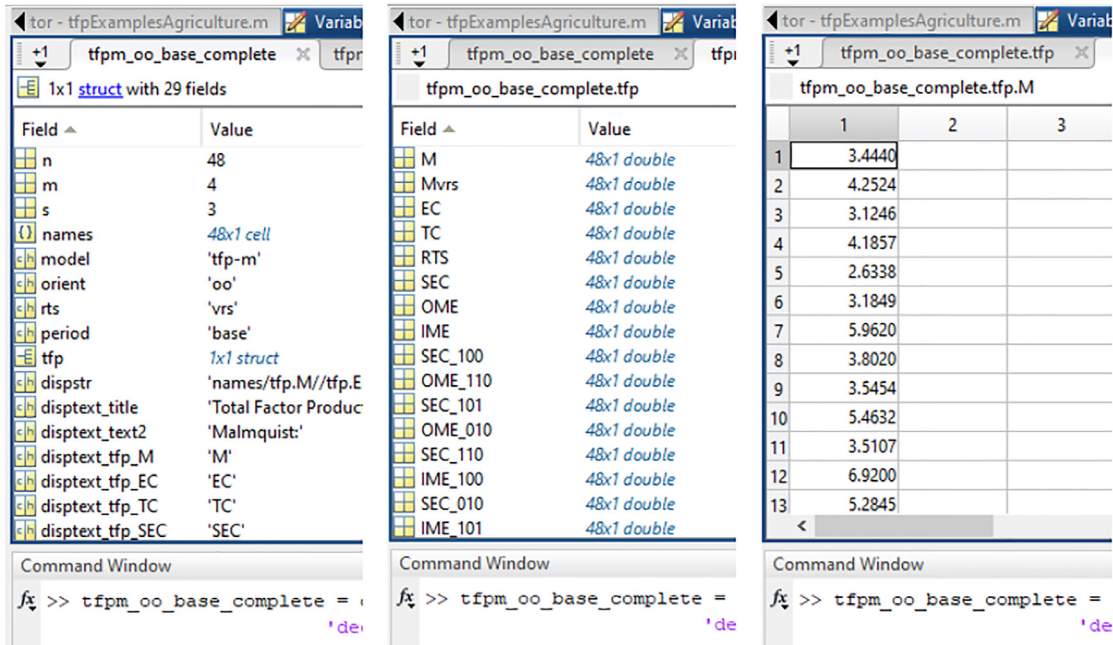


Fig. 1. Visualization of the structures for model parameters and results.

MPI requires calculation of the output- or input-orientated distance of observation  $(\mathbf{x}^{kt}, \mathbf{y}^{kt})$  in two consecutive periods (say, base period  $t = 0$  and comparison period  $t = 1$ ) to the frontier of a certain benchmark technology exhibiting CRS. The imposition of CRS is necessary for the distance function to satisfy the homogeneity conditions that ensure that the MPI has the required proportionality properties, see Balk and Zofio (2018). Thus in general the benchmark technology will differ from any actual technology.

It is usual to take the cone technology of a certain period  $t$ ,  $\tilde{S}^t$ , as benchmark. Its output distance function is defined by  $\tilde{D}_o^t(\mathbf{x}, \mathbf{y}) \equiv \inf\{\delta \mid \delta > 0, (\mathbf{x}, \mathbf{y}/\delta) \in \tilde{S}^t\}$ . Operationally, within the DEA framework, this can be calculated by solving the program  $\tilde{D}_o^t(\mathbf{x}, \mathbf{y})^{-1} = \max_{\phi, \lambda} \{\phi \mid \mathbf{x} \geq X^t \lambda, \mathbf{y} \phi \leq Y^t \lambda, \lambda \geq \mathbf{0}\}$ . Then  $(\mathbf{x}, \mathbf{y}/\tilde{D}_o^t(\mathbf{x}, \mathbf{y}))$  is the point on the frontier of the period  $t$  cone technology that is obtained by holding the input quantity vector  $\mathbf{x}$  constant while radially expanding the output quantity vector  $\mathbf{y}$ .

The input distance function is defined as  $\tilde{D}_i^t(\mathbf{x}, \mathbf{y}) \equiv \sup\{\delta \mid \delta > 0, (\mathbf{x}/\delta, \mathbf{y}) \in \tilde{S}^t\}$ , and can be calculated by solving the program  $\tilde{D}_i^t(\mathbf{x}, \mathbf{y})^{-1} = \min_{\theta, \lambda} \{\theta \mid \mathbf{x}\theta \geq X^t \lambda, \mathbf{y} \leq Y^t \lambda, \lambda \geq \mathbf{0}\}$ . Then  $(\mathbf{x}/\tilde{D}_i^t(\mathbf{x}, \mathbf{y}), \mathbf{y})$  is the point on the frontier of the period  $t$  cone technology that is obtained by holding the output quantity vector  $\mathbf{y}$  constant while radially contracting the input quantity vector  $\mathbf{x}$ . Notice that  $\tilde{D}_i^t(\mathbf{x}, \mathbf{y}) = 1/\tilde{D}_o^t(\mathbf{x}, \mathbf{y})$ .

The counterparts of the above output and input distance functions, defined on the actual technology  $S^t$  which in general exhibits VRS, denoted by  $D_o^t(\mathbf{x}, \mathbf{y})$  and  $D_i^t(\mathbf{x}, \mathbf{y})$ , are computed in the same way, with  $\mathbf{e}\lambda = \mathbf{1}$  as additional constraint.

### 3.1. The output-orientated MPI

The output-orientated MPI, conditional on the period  $t$  cone technology, for a certain DMU, is defined by<sup>4</sup>

$$\tilde{M}_o^t(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \equiv \frac{\tilde{D}_o^t(\mathbf{x}^1, \mathbf{y}^1)}{\tilde{D}_o^t(\mathbf{x}^0, \mathbf{y}^0)} \quad (1)$$

<sup>4</sup> Here and in the sequel the superscript  $k$ , designating a specific DMU, is deleted to simplify presentation.

Selecting the base period cone technology then leads to  $\tilde{M}_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)$ , and selecting the comparison period cone technology leads to  $\tilde{M}_o^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)$ . The TFP toolbox calculates both, as well as their geometric mean. Let us start with the first option.

#### 3.1.1. The base-period-output-orientated MPI

Following Balk and Zofio (2018), who provide meaningful theoretical interpretations for the different factors, the first extended decomposition of the base-period-output-orientated MPI (termed ‘Path A’) is

$$\tilde{M}_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_o^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0) \times OME^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{y}^0) \quad (2)$$

In this expression there are four mutually independent factors, with the following interpretation:

- Efficiency change:  $EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = D_o^1(\mathbf{x}^1, \mathbf{y}^1)/D_o^0(\mathbf{x}^0, \mathbf{y}^0)$ , representing the change in the technical efficiency of the DMU, also known as the *catch-up effect*.
- Technological change:  $TC_o^{1,0}(\mathbf{x}^1, \mathbf{y}^1) = D_o^0(\mathbf{x}^1, \mathbf{y}^1)/D_o^1(\mathbf{x}^1, \mathbf{y}^1)$ , capturing the change in the actual technological frontier, also known as the *frontier-shift effect*.
- Radial scale and input mix effect:  $SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0) = [\tilde{D}_o^0(\mathbf{x}^1, \mathbf{y}^0)/D_o^0(\mathbf{x}^1, \mathbf{y}^0)] \times [D_o^0(\mathbf{x}^0, \mathbf{y}^0)/\tilde{D}_o^0(\mathbf{x}^0, \mathbf{y}^0)]$ , corresponding to scale efficiency improvements associated to radial increases in the input quantities, and the additional effect from changes in the input quantity mix.<sup>5</sup>
- Output mix effect:  $OME^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{y}^0) = [\tilde{D}_o^0(\mathbf{x}^1, \mathbf{y}^1)/D_o^0(\mathbf{x}^1, \mathbf{y}^1)] \times [D_o^0(\mathbf{x}^1, \mathbf{y}^0)/\tilde{D}_o^0(\mathbf{x}^1, \mathbf{y}^0)]$ , showing the counterpart effect associated to changes in the output quantity mix.

Factors with values greater than 1 contribute to productivity growth (e.g., through technical efficiency gains or technological progress), while factors whose values are smaller than 1 are detrimental.

<sup>5</sup> Balk and Zofio (2018) show that  $SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0)$  can be decomposed into a radial scale effect and an input quantity mix effect. The implementation in a DEA framework is however too complicated to be practically useful.

An alternative decomposition of the base-period-output-orientated MPI reverses the order in which changes in the input and output space take place in the last two factors in expression (2). This yields

$$\tilde{M}_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_o^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^1) \times OME^0(\mathbf{x}^0, \mathbf{y}^1, \mathbf{y}^0), \tag{3}$$

which is termed ‘Path B’. The differences between this decomposition and that in expression (2) are subtle but noteworthy. The parts capturing efficiency change and technological change are identical. The factor capturing the radial scale effect and the input mix effect is conditional on  $\mathbf{y}^0$  in expression (2), but on  $\mathbf{y}^1$  in expression (3). The reverse happens with the output mix effect; this effect is conditional on  $\mathbf{x}^1$  in expression (2), but on  $\mathbf{x}^0$  in expression (3).

Consequently, there are two, equally meaningful, decompositions of the Malmquist productivity index  $\tilde{M}_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)$ . If there is no preference for one of them then the geometric mean of expressions (2) and (3) can be taken, reading

$$\tilde{M}_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_o^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times [SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0) \times SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^1)]^{1/2} \times [OME^0(\mathbf{x}^0, \mathbf{y}^1, \mathbf{y}^0) \times OME^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{y}^0)]^{1/2}. \tag{4}$$

To compute the MPI in MATLAB the user calls the `deatfpm(X, Y, ...)` function with i) the `orient` parameter set to the output orientation `oo`; 2) the `period` parameter set to `base`; and iii) the `decomposition` parameter `decomp` set to `complete`. With the optional parameter `names` we can specify a cell string with the names of the DMUs, which in this case are the names of the U.S. states.<sup>6</sup>

The model parameters and all the ancillary information can be found in the editor under `tfpm_oo_base_complete`. The results are displayed in the ‘command window’, and the structure with all the variables and results can be found in the workspace by clicking on `tfpm_oo_base_complete.tfpm`. Fig. 1 shows the structures for the model parameters and results in the MATLAB editor.

The output of the function successively shows the MPI (M), the technical efficiency change factor (EC), the technological change factor (TC), the geometric mean of the scale effect factors (SEC), the geometric mean of the output mix effect factors (OME), and the separate values of the two last factors corresponding to ‘Path A’ and ‘Path B’ above. These values are identified by the time superscripts of the input and output arguments in each factor; for example, `SEC_100` corresponds to  $SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0)$  in expressions (2) and (4) above. The function also returns the main descriptive statistics of the various factors. Finally, a relevant clarification concerns the nature of returns to scale. The Malmquist index is defined with respect to a cone technology, independently of whether the actual technology is characterized by global CRS or not. Thus the default decomposition is based on VRS technologies, which is identified by the corresponding returns-to-scale header below; i.e., “Returns to scale: `vrs` (Variable)”. The case of a technology characterized by global CRS is considered in the following subsection.

```
> tfpm_oo_base_complete = deatfpm(X, Y, 'orient', 'oo', 'period', 'base', ...
' decomp', 'complete', 'names', STATE_NAME);
> tfpdisp(tfpm_oo_base_complete)
```

```
Total Factor Productivity (TFP)
DMUs: 48      Inputs: 4      Outputs: 3
Model: tfp-m
Orientation: oo (Output orientated)
Returns to scale: vrs (Variable)
Malmquist: Reference period is base period
```

DMU	M	EC	TC	SEC	OME	SEC_100	OME_110	SEC_101	OME_010
AL	3.4440	1.3105	2.6698	0.9843	1.0000	0.9843	1.0000	0.9843	1.0000
AR	4.2524	1.3527	3.2174	0.9868	0.9902	0.9949	0.9821	0.9786	0.9984
AZ	3.1246	0.9722	3.5293	0.9297	0.9796	0.9491	0.9595	0.9107	1.0000
...									
WI	1.9343	1.0397	1.8528	1.0318	0.9732	1.0172	0.9872	1.0466	0.9595
WV	0.8177	1.0020	0.8737	0.9715	0.9614	1.0114	0.9235	0.9332	1.0009
WY	1.1263	0.8509	1.4799	0.9370	0.9545	0.9634	0.9284	0.9113	0.9814
Mean	3.1599	1.1101	2.9944	0.9686	0.9793	0.9716	0.9765	0.9662	0.9598
Std	1.3456	0.1823	1.2485	0.0555	0.0352	0.0613	0.0295	0.0582	0.1142
Max	6.9200	1.4695	6.6537	1.0318	1.0100	1.0316	1.0202	1.0684	1.0115
Min	0.8177	0.7703	0.8737	0.6977	0.7776	0.6714	0.8742	0.7252	0.4022

M = Malmquist. EC = Efficiency Change. TC = Technological Change.  
SEC = Scale Effect. OME = Output Mix Effect.

We exemplify individual results by discussing the productivity trends of the first and last three states. Most states increased their agricultural productivity, Illinois (IL) leading with a 592.00% increase, corresponding to the percentage change of the maximum MPI value. Exceptions exhibiting productivity decline are Oklahoma (OK), unreported, and West Virginia (WV), with a 18.23% reduction (the minimum value of the MPI). The main component of productivity growth, on average 215.99%, is technological progress, which generally outpaces efficiency gains by far (on average they contribute 199.44% and 11.01%, respectively). Changes in the scale of production or the output

<sup>6</sup> If the optional parameter `names` is omitted, the DMUs will be numbered from 1 to *n*.

mix do not play a significant role, implying that there have no been relevant changes in the production scale and the input and output structure of the states over the forty-four year period 1960–2004. Similar comments can be made for each individual state.<sup>7</sup>

### 3.1.2. Related decompositions of the base-period-output-orientated MPI

The above four-factor decompositions can be related to simpler proposals in the literature. First, they generalize the earlier proposal by Ray and Desli (1997). By merging the scale effect and the output mix effect in expressions (2) or (3), both decompositions reduce to the same three-factor decomposition:

$$\tilde{M}_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_o^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times RTS^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0), \quad (5)$$

where  $RTS^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \equiv [\tilde{D}_o^0(\mathbf{x}^1, \mathbf{y}^1)/D_o^0(\mathbf{x}^1, \mathbf{y}^1)]/[\tilde{D}_o^0(\mathbf{x}^0, \mathbf{y}^0)/D_o^0(\mathbf{x}^0, \mathbf{y}^0)]$ . Lovell (2003) suggests that this factor measures the contribution of returns-to-scale to productivity change. Following this interpretation, if it is greater than 1, increasing returns to scale result in productivity growth, while if it is smaller than 1, decreasing returns to scale are detrimental. If equal to one, constant returns to scale prevails at the observed input-output values.

To obtain this simpler decomposition, the syntax is the same as in the extended one but with the parameter `decomp` set to `rts`. In the output table the column `RTS` contains the contribution of returns-to-scale to productivity change.

```
> tfpm_oo_base_rts = deatfpm(X, Y, 'orient', 'oo', 'period', 'base', ...
' decomp', 'rts', 'names', STATE_NAME);
> tfpdisp(tfpm_oo_base_rts)
```

```
-----
Total Factor Productivity (TFP)
DMUs: 48      Inputs: 4      Outputs: 3
Model: tfp-m
Orientation: oo (Output orientated)
Returns to scale: vrs (Variable)
Malmquist: Reference period is base period
-----
DMU|      M|      EC|      TC|      RTS|
-----
AL| 3.4440| 1.3105| 2.6698| 0.9843|
AR| 4.2524| 1.3527| 3.2174| 0.9771|
AZ| 3.1246| 0.9722| 3.5293| 0.9107|
...
WI| 1.9343| 1.0397| 1.8528| 1.0041|
WV| 0.8177| 1.0020| 0.8737| 0.9340|
WY| 1.1263| 0.8509| 1.4799| 0.8944|
-----
Mean| 3.1599| 1.1101| 2.9944| 0.9489|
Std| 1.3456| 0.1823| 1.2485| 0.0693|
Max| 6.9200| 1.4695| 6.6537| 1.0316|
Min| 0.8177| 0.7703| 0.8737| 0.6808|
-----
```

Notice that the contribution of returns-to-scale to productivity growth, consistent with small values of the scale and output mix effects, is quite limited. In the case of West Virginia (WV) the existence of decreasing returns to scale, along with the subsidence of the production frontier, associated to values of technological change smaller than one, results in the previously reported productivity decline to the tune of –18.23%.

Second, the technical efficiency change and technological change terms common to expressions (2), (3), and (4) correspond to the base-period-output-orientated index proposed by Caves et al. (1982) (CCD),

$$M_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \equiv EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_o^{1,0}(\mathbf{x}^1, \mathbf{y}^1) = \frac{D_o^0(\mathbf{x}^1, \mathbf{y}^1)}{D_o^0(\mathbf{x}^0, \mathbf{y}^0)}. \quad (6)$$

Substituting expression (6) in any of the above decompositions yields alternative expressions,

$$\tilde{M}_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = M_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0) \times OME^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{y}^0) \quad (7)$$

$$\tilde{M}_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = M_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^1) \times OME^0(\mathbf{x}^0, \mathbf{y}^1, \mathbf{y}^0) \quad (8)$$

$$\tilde{M}_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = M_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times [SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0) \times SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^1)]^{1/2} \times [OME^0(\mathbf{x}^0, \mathbf{y}^1, \mathbf{y}^0) \times OME^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{y}^0)]^{1/2}. \quad (9)$$

To obtain this decomposition the parameter `decomp` must be set to `ccd`.

<sup>7</sup> Descriptive statistics, productivity distributions, and associated graphs, can be easily calculated and plotted in MATLAB.

```
> tfpm_oo_base_ccd = deatfpm(X, Y, 'orient', 'oo', 'period', 'base', ...
'decom', 'ccd', 'names', STATE_NAME);
> tfpdisp(tfpm_oo_base_ccd)
```

```
-----
Total Factor Productivity (TFP)
DMUs: 48      Inputs: 4      Outputs: 3
Model: tfp-m
Orientation: oo (Output orientated)
Returns to scale: vrs (Variable)
Malmquist: Reference period is base period
-----
```

DMU	M	MCCD	SEC	OME	SEC_100	OME_110	SEC_101	OME_010
AL	3.4440	3.4989	0.9843	1.0000	0.9843	1.0000	0.9843	1.0000
AR	4.2524	4.3521	0.9868	0.9902	0.9949	0.9821	0.9786	0.9984
AZ	3.1246	3.4310	0.9297	0.9796	0.9491	0.9595	0.9107	1.0000
...								
WI	1.9343	1.9263	1.0318	0.9732	1.0172	0.9872	1.0466	0.9595
WV	0.8177	0.8755	0.9715	0.9614	1.0114	0.9235	0.9332	1.0009
WY	1.1263	1.2593	0.9370	0.9545	0.9634	0.9284	0.9113	0.9814
Mean	3.1599	3.3038	0.9686	0.9793	0.9716	0.9765	0.9662	0.9598
Std	1.3456	1.3485	0.0555	0.0352	0.0613	0.0295	0.0582	0.1142
Max	6.9200	7.1328	1.0318	1.0100	1.0316	1.0202	1.0684	1.0115
Min	0.8177	0.8755	0.6977	0.7776	0.6714	0.8742	0.7252	0.4022

```
-----
```

In the extant literature on productivity measurement, the CCD index frequently figures under the name ‘the (output orientated) Malmquist productivity index’. However, the CCD index, expression (6), cannot be regarded as a productivity index. First, it cannot be written as a measure of aggregate output growth divided by a measure of aggregate input growth, the corresponding aggregators being non-negative non-decreasing linearly homogenous scalar functions. O’Donnell (2012) uses the term “multiplicatively complete” to refer to TFP indexes constructed in this way. More importantly, the CCD index does not satisfy the proportionality property, as initially shown by Grifell-Tatjé and Lovell (1995), unless the actual base period technology exhibits CRS. In the specific case of a technology satisfying CRS, all the SEC and OME terms in expressions (7), (8), and (9) become identical to 1, and the MPI  $\check{M}_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)$  reduces to the CCD index  $M_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)$ . The MPI is then decomposed as initially proposed by Färe et al. (1994),<sup>8</sup>

$$\check{M}_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = \check{E}C_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times \check{T}C_o^{1,0}(\mathbf{x}^1, \mathbf{y}^1). \quad (10)$$

The two-factor decomposition of the MPI for a CRS technology can be calculated by setting the parameter `decomp` to `crs`,

```
> tfpm_oo_base_crs = deatfpm(X, Y, 'orient', 'oo', 'period', 'base', ...
'decom', 'crs', 'names', STATE_NAME);
> tfpdisp(tfpm_oo_base_crs)
```

<sup>8</sup> The toolbox offers the possibility of testing the hypothesis that a certain technology is characterized by VRS or CRS. The function implements the algorithms developed by Simar and Wilson (2002), following Bogetoft and Otto (2011), to determine whether the VRS and CRS distance function values are significantly different or not. The test can be performed by calling the `deatestrtsm(X, Y, ...)` function, with the `orient` parameter set to the same orientation as the MPI.

```

-----
Total Factor Productivity (TFP)
DMUs: 48      Inputs: 4      Outputs: 3
Model: tfp-m
Orientation: oo (Output orientated)
Returns to scale: crs (Constant)
Malmquist: Reference period is base period
-----

```

DMU	M	EC	TC
AL	3.4440	1.2297	2.8006
AR	4.2524	1.3547	3.1389
AZ	3.1246	0.9698	3.2218
...			
WI	1.9343	0.8946	2.1620
WV	0.8177	0.9980	0.8193
WY	1.1263	0.8680	1.2975
-----			
Mean	3.1599	1.1030	2.8994
Std	1.3456	0.1901	1.2623
Max	6.9200	1.5653	6.4489
Min	0.8177	0.7161	0.8193
-----			

Notice that, as the MPI does not depend on an actual technology, its outcomes are identical to those in the previous tables. The decompositions in EC and TC factors, however, are different. The differences are not large, as for most of the DMUs the RTS factor in expression (5) resides in the neighbourhood of 1. We again see the prominent role played by technological change in most of the American states, except Oklahoma (OK), not reported in the table above, and West Virginia (WV).

### 3.1.3. The comparison-period-output-orientated MPI

As anticipated, we can also select the comparison period technology,  $\tilde{S}^1$ , as benchmark. The comparison-period-output-orientated MPI is then given by  $\tilde{M}_o^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)$ . The complete four-factor decomposition analogous to expressions (2) and (3) are, respectively,

$$\tilde{M}_o^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_o^{1,0}(\mathbf{x}^0, \mathbf{y}^0) \times SEC_o^1(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0) \times OME^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{y}^0), \quad (11)$$

and

$$\tilde{M}_o^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_o^{1,0}(\mathbf{x}^0, \mathbf{y}^0) \times SEC_o^1(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^1) \times OME^1(\mathbf{x}^0, \mathbf{y}^1, \mathbf{y}^0), \quad (12)$$

which are identified by 'Path C' and 'Path D' in Balk and Zofio (2018). Again, if there is no preference for either of the two then it is advised to take the geometric mean of these two decompositions,

$$\tilde{M}_o^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_o^{1,0}(\mathbf{x}^0, \mathbf{y}^0) \times [SEC_o^1(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0) \times SEC_o^1(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^1)]^{1/2} \times [OME^1(\mathbf{x}^0, \mathbf{y}^1, \mathbf{y}^0) \times OME^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{y}^0)]^{1/2}. \quad (13)$$

The comparison-period-output-orientated CCD index is defined by

$$M_o^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \equiv EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_o^{1,0}(\mathbf{x}^0, \mathbf{y}^0) = \frac{D_o^1(\mathbf{x}^1, \mathbf{y}^1)}{D_o^1(\mathbf{x}^0, \mathbf{y}^0)}. \quad (14)$$

The complete decomposition of the comparison-period-output-orientated MPI is calculated by calling the `deatfpm(X, Y, ...)` function, replacing `base` by `comparison` in the period parameter, and leaving the other parameters unchanged:

```

> tfpm_oo_comp_complete = deatfpm(X, Y, 'orient', 'oo', 'period', 'comparison', ...
'decomp', 'complete', 'names', STATE_NAME);
> tfpdisp(tfpm_oo_comp_complete)

```

The output has the same structure as the base period analogue and is therefore not reported here. Also, obtaining the related decompositions is done by setting the parameter `decomp` to either `rts`, `ccd`, or `crs`.

### 3.1.4. The geometric-mean-output-orientated MPI

In general the output-orientated MPIs with base and comparison period benchmarks will deliver different results. A compromise between the two MPIs is their geometric mean,

$$\tilde{M}_o(x^1, y^1, x^0, y^0) \equiv [\tilde{M}_o^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times \tilde{M}_o^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)]^{1/2} = \left[ \frac{\tilde{D}_o^0(\mathbf{x}^1, \mathbf{y}^1) \tilde{D}_o^1(\mathbf{x}^1, \mathbf{y}^1)}{\tilde{D}_o^0(\mathbf{x}^0, \mathbf{y}^0) \tilde{D}_o^1(\mathbf{x}^0, \mathbf{y}^0)} \right]^{1/2}. \quad (15)$$



This index can be decomposed into the following factors,

$$\begin{aligned} \check{M}_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) &= EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times [TC_o^{1,0}(\mathbf{x}^0, \mathbf{y}^0)TC_o^{1,0}(\mathbf{x}^1, \mathbf{y}^1)]^{1/2} \\ &\times [SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0)SEC_o^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^1)SEC_o^1(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0)SEC_o^1(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^1)]^{1/4} \\ &\times [OME^0(\mathbf{x}^0, \mathbf{y}^1, \mathbf{y}^0)OME^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{y}^0)OME^1(\mathbf{x}^0, \mathbf{y}^1, \mathbf{y}^0)OME^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{y}^0)]^{1/4}. \end{aligned} \tag{16}$$

This decomposition is computed by setting the period parameter to `geomean` in the `deatfpm(X, Y, ...)` function.

```
> tfpm_oo_geomean_complete = deatfpm(X, Y, 'orient', 'oo', 'period', 'geomean', ...
'decom', 'complete', 'names', STATE_NAME);
> tfpdisp(tfpm_oo_geomean_complete)
```

```
-----
Total Factor Productivity (TFP)
DMUs: 48      Inputs: 4      Outputs: 3
Model: tfp-m
Orientation: oo (Output orientated)
Returns to scale: vrs (Variable)
Malmquist: Reference period is Geometric mean
-----
```

DMU	M	EC	TC	SEC	OME
AL	1.6964	1.3105	1.2565	1.0018	1.0283
AR	1.8945	1.3527	1.3553	1.0170	1.0161
AZ	1.4400	0.9722	1.5511	0.9667	0.9878
...					
WI	1.2578	1.0397	1.1133	1.0363	1.0485
WV	0.7457	1.0020	0.7787	0.9829	0.9723
WY	0.8164	0.8509	1.0042	0.9801	0.9748
Mean	1.5494	1.1101	1.3352	0.9946	1.0318
Std	0.4395	0.1823	0.3016	0.0366	0.0544
Max	2.6882	1.4695	2.0343	1.0869	1.2041
Min	0.5219	0.7703	0.6472	0.8680	0.9507

```
-----
```

The corresponding individual terms associated to 'Path A', 'Path B', 'Path C', or 'Path D' can be recovered by running the previous base or comparison options. It is then possible to confirm that the two decompositions are indeed different and their geometric mean is a useful compromise. Also, the related decompositions are obtained by setting the parameter `decomp` to `rts`, `ccd`, or `crs`.

The importance of the benchmark period is clear in the case of US agriculture. The geometric mean of the base and comparison period benchmarks delivers a 54.94% increase in productivity over the 44 year period. This is substantially lower than the previously reported result from the base period benchmark, 215.99%, suggesting that the comparison period benchmark delivers a rather limited productivity change. Since the EC factor is common to both decompositions (11.01%), the reduction in productivity growth corresponds to lower TC to the tune of 33.52%. Finally, regardless of the choice of the benchmark technology, input and output scale and mix effects barely contribute to productivity growth, their magnitudes being close to one.

### 3.2. The input-orientated MPI

Recall that a cone technology exhibits CRS, thus  $\check{D}_o^t(\mathbf{x}^t, \mathbf{y}^t) = 1/\check{D}_i^t(\mathbf{x}^t, \mathbf{y}^t)$ . Thus the output-orientated MPI defined by expression (1) can also be written as

$$\check{M}_o^t(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) = \frac{\check{D}_i^t(\mathbf{x}^0, \mathbf{y}^0)}{\check{D}_i^t(\mathbf{x}^1, \mathbf{y}^1)} \equiv \check{M}_i^t(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0), \tag{17}$$

that is, as input-orientated MPI, conditional on the period  $t$  cone technology,  $\check{S}^t$ .

The options for decomposing an input-orientated MPI run parallel to those already presented for the output-orientated counterpart, and therefore it is sufficient to describe the changes that need to be introduced in the MATLAB functions to obtain the different decompositions. For formal definitions and a graphical representation of the input decompositions see Balk and Zofio (2018, Section 4). In general, all that is required is to substitute `io` for `oo` in the `orient` parameter in the MPI function.

The base-period-input-orientated MPI, defined as  $\check{M}_i^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)$ , is calculated and decomposed by running the following code:

```
> tfpm_io_base_complete = deatfpm(X, Y, 'orient', 'io', 'period', 'base', ...
'decom', 'complete', 'names', STATE_NAME)
```

Following procedures identical to those for the output-orientated MPI the related decompositions can be obtained. Merging the input mix and scale effects into the `returns-to-scale` factor yields the three-factor decomposition counterpart to expression (5). To obtain this simpler decomposition the parameter `decomp` must be changed to `rts`.

```
> tfpm_io_base_rts = deatfpm(X, Y, 'orient', 'io', 'period', 'base', ...
'decom', 'rts', 'names', STATE_NAME)
```

As in expression (6), one can merge the technical efficiency change and technological change factors into the base-period-input-orientated version of CCD index. For this decomposition the parameter `decomp` must be changed from `complete` to `ccd`.

```
> tfpm_io_base_ccd = deatfpm(X, Y, 'orient', 'io', 'period', 'base', ...
'decom', 'ccd', 'names', STATE_NAME)
```

The CRS version can be calculated by setting the parameter `decomp` to `crs`.

```
> tfpm_io_base_crs = deatfpm(X, Y, 'orient', 'io', 'period', 'base', ...
'decom', 'crs', 'names', STATE_NAME)
```

Choosing the cone technology  $\tilde{S}^1$  as benchmark leads to the comparison-period-input-orientated MPI,  $\tilde{M}_i^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)$ . The complete decomposition requires changing base to comparison in the period parameter of the `deatfpm(X, Y, ...)` function.

```
> tfpm_io_comp_complete = deatfpm(X, Y, 'orient', 'io', 'period', 'comparison', ...
'decom', 'complete', 'names', STATE_NAME)
```

Finally, the geometric mean of the base and comparison period perspectives,  $\tilde{M}_i(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)$   $\equiv [M_i^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times M_i^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)]^{1/2}$ , as in the case of the output-orientated MPI (expression (15)) is obtained by setting the period parameter to `geomean`.

```
> tfpm_io_geomean_complete = deatfpm(X, Y, 'orient', 'io', 'period', 'geomean', ...
'decom', 'complete', 'names', STATE_NAME)
```

Related decompositions for the comparison-period- and geometric-mean-input-orientated MPI are obtained by setting the parameter `decomp` to either `rts`, `ccd`, or `crs`. For each variant, the output of the function is interpreted in the same way as their base-period counterparts.

#### 4. Moorsteen-Bjurek productivity index

The second definition of a productivity index based only on quantities takes into account both the output and input orientations. Specifically, the family of Moorsteen-Bjurek productivity indices (MBPI) is defined as the ratio of a Malmquist output quantity index to a Malmquist input quantity index, conditional on a benchmark technology  $S^t$ . A Malmquist output quantity index, comparing output quantities  $\mathbf{y}^1$  to  $\mathbf{y}^0$ , conditional on certain input quantities  $\bar{\mathbf{x}}$ , is defined as  $Q_o^t(\mathbf{y}^1, \mathbf{y}^0, \bar{\mathbf{x}}) \equiv D_o^t(\bar{\mathbf{x}}, \mathbf{y}^1) / D_o^t(\bar{\mathbf{x}}, \mathbf{y}^0)$ . Similarly, a Malmquist input quantity index, comparing input quantities  $\mathbf{x}^1$  to  $\mathbf{x}^0$ , conditional on certain output quantities  $\bar{\mathbf{y}}$ , is defined as  $Q_i^t(\mathbf{x}^1, \mathbf{x}^0, \bar{\mathbf{y}}) \equiv D_i^t(\mathbf{x}^1, \bar{\mathbf{y}}) / D_i^t(\mathbf{x}^0, \bar{\mathbf{y}})$ . Both indices can be traced back to suggestions by Moorsteen (1961), and their properties were extensively discussed in Balk (1998). Typically, in empirical applications involving many DMUs,  $\bar{\mathbf{x}}$  and  $\bar{\mathbf{y}}$  would be chosen as vectors of sample means. This is the approach followed in the toolbox. The MBPI is then defined by

$$MB^t(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}}) \equiv \frac{Q_o^t(\mathbf{y}^1, \mathbf{y}^0, \bar{\mathbf{x}})}{Q_i^t(\mathbf{x}^1, \mathbf{x}^0, \bar{\mathbf{y}})} = \frac{D_o^t(\bar{\mathbf{x}}, \mathbf{y}^1) / D_o^t(\bar{\mathbf{x}}, \mathbf{y}^0)}{D_i^t(\mathbf{x}^1, \bar{\mathbf{y}}) / D_i^t(\mathbf{x}^0, \bar{\mathbf{y}})} \quad (18)$$

The last term shows that the MBPI can be seen as a ratio of two productivity levels. Up to a scalar normalization, and conditional on  $\bar{\mathbf{x}}$  and  $\bar{\mathbf{y}}$ , the productivity level at the input-output situation  $(\mathbf{x}, \mathbf{y})$  is thereby measured as  $D_o^t(\bar{\mathbf{x}}, \mathbf{y}) / D_i^t(\mathbf{x}, \bar{\mathbf{y}})$ , where superscript  $t$  refers to the benchmark technology  $S^t$ . Thus, the MBPI belongs to the class of “multiplicatively complete” TFP indices, as defined by O'Donnell (2012).

Based on the properties of the MBPI, Balk and Zofio (2018) show that this index can be decomposed into factors corresponding to those already shown and interpreted. Decompositions can be based on output distance functions or input distance functions. As benchmark the technologies of the base period 0 and comparison 1 are used. We follow the order in which the MPI decompositions were discussed.

##### 4.1. The output-orientated decomposition of MBPI

###### 4.1.1. The base-period MBPI

Taking as benchmark the base-period technology, along ‘Path A’ the MBPI is decomposed as

$$MB^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}}) = EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_o^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times \left( \frac{D_o^0(\mathbf{x}^0, \mathbf{y}^0) D_i^0(\mathbf{x}^0, \bar{\mathbf{y}})}{D_o^0(\mathbf{x}^1, \mathbf{y}^0) D_i^0(\mathbf{x}^1, \bar{\mathbf{y}})} \right) \times \left( \frac{D_o^0(\bar{\mathbf{x}}, \mathbf{y}^1) D_o^0(\mathbf{x}^1, \mathbf{y}^0)}{D_o^0(\bar{\mathbf{x}}, \mathbf{y}^0) D_o^0(\mathbf{x}^1, \mathbf{y}^1)} \right) \quad (19)$$

The factors on the right-hand side of the equality sign represent, respectively, technical efficiency change, technological change (conditional on  $(\mathbf{x}^1, \mathbf{y}^1)$ ), the radial scale and input mix effect (conditional on  $\mathbf{y}^0$ ), and the output mix effect (conditional on  $\mathbf{x}^1$ ).

With the same benchmark technology, along ‘Path B’ the following decomposition is obtained:

$$MB^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}}) = EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_o^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times \left( \frac{D_o^0(\mathbf{x}^0, \mathbf{y}^1) D_i^0(\mathbf{x}^0, \bar{\mathbf{y}})}{D_o^0(\mathbf{x}^1, \mathbf{y}^1) D_i^0(\mathbf{x}^1, \bar{\mathbf{y}})} \right) \times \left( \frac{D_o^0(\bar{\mathbf{x}}, \mathbf{y}^1) D_o^0(\mathbf{x}^0, \mathbf{y}^0)}{D_o^0(\bar{\mathbf{x}}, \mathbf{y}^0) D_o^0(\mathbf{x}^0, \mathbf{y}^1)} \right) \quad (20)$$

The first two terms are the same but the radial scale and input mix effect are now conditional on  $\mathbf{y}^1$ , while the output mix effect is conditional on  $\mathbf{x}^0$ . The geometric mean of the two decompositions is computed by the `deatfpm(X, Y, ...)` function with: (i) the

orient parameter set to output orientation oo; (ii) the period parameter set to base; and (iii) the decomposition parameter decomp set to complete. In this function, the (arithmetic) averages of the input quantities  $\bar{x}$  and output quantities  $\bar{y}$  in the two periods 0 and 1 are automatically calculated from the arrays of X and Y.

```
> tfpmb_oo_base_complete = deatfpmb(X, Y, 'orient', 'oo', 'period', 'base', ...
' decomp', 'complete', 'names', STATE_NAME);
> tfpdisp(tfpmb_oo_base_complete)
```

-----  
Total Factor Productivity (TFP)

DMUs: 48  
Inputs: 4      Outputs: 3  
Model: tfp-mb  
Orientation: oo (Output orientated)  
Returns to scale: vrs (Variable)

Moorsteen--Bjurek index:  
Reference period is base period

DMU	MB	EC	TC	SEC	OME	SEC_100	OME_110	SEC_101	OME_010
AL	3.4377	1.3105	2.6698	0.9825	1.0000	0.9825	1.0000	0.9825	1.0000
AR	5.1654	1.3527	3.2174	1.0711	1.1081	1.1092	1.0701	1.0344	1.1474
AZ	2.7241	0.9722	3.5293	0.7914	1.0032	0.7827	1.0143	0.8002	0.9922
...									
WI	2.3098	1.0397	1.8528	1.0245	1.1704	0.9758	1.2288	1.0756	1.1149
WV	0.8238	1.0020	0.8737	0.9874	0.9529	1.0289	0.9146	0.9477	0.9929
WY	1.2377	0.8509	1.4799	0.9699	1.0133	1.0260	0.9579	0.9169	1.0719

The output of the function successively shows the MBPI (MB), EC, TC, and the geometric mean of the SEC and OME factors. The last four columns show the separate values of the SEC and OME factors corresponding to 'Path A' and 'Path B', identified by the time superscripts of each of their input and output arguments. For instance, SEC\_100 identifies the third factor in expression (19), counterpart to  $SEC_0^0(x^1, x^0, y^0)$  in expression (3).

Regarding US agricultural productivity growth, it is interesting to compare the results corresponding to the Malmquist and Moorsteen–Bjurek definitions. The difference between the two is rather small, as average growth for the latter amounts to 229.52% versus 215.99% for the former. Since the efficiency change and technological change factors are common to both indices, the difference between the two must be due to the scale and output mix effects. While these effects were clearly negative in the case of the MPI, with values well below unity, here we observe that they positively contribute to productivity growth. Nevertheless we conclude that, for the US agriculture, the choice of the Malmquist or Moorsteen–Bjurek definition does not result in relevant differences at the aggregate or individual level. Indeed, the Spearman rank correlation between the two indices is statistically significant at the 1% confidence level with a value of  $\rho = 0.8629$ .

4.1.2. Related decompositions

The MBPIs in expressions (19) and (20) are related to the CCD index, defined in expression (6). Thus the two expressions can be rewritten as

$$MB^0(x^1, y^1, x^0, y^0; \bar{x}, \bar{y}) = M_0^0(x^1, y^1, x^0, y^0) \times \left( \frac{D_0^0(x^0, y^0) D_i^0(x^0, \bar{y})}{D_0^0(x^1, y^0) D_i^0(x^1, \bar{y})} \right) \times \left( \frac{D_0^0(\bar{x}, y^1) D_0^0(x^1, y^0)}{D_0^0(\bar{x}, y^0) D_0^0(x^1, y^1)} \right) \tag{21}$$

$$MB^0(x^1, y^1, x^0, y^0; \bar{x}, \bar{y}) = M_0^0(x^1, y^1, x^0, y^0) \times \left( \frac{D_0^0(x^0, y^1) D_i^0(x^0, \bar{y})}{D_0^0(x^1, y^1) D_i^0(x^1, \bar{y})} \right) \times \left( \frac{D_0^0(\bar{x}, y^1) D_0^0(x^0, y^0)}{D_0^0(\bar{x}, y^0) D_0^0(x^0, y^1)} \right) \tag{22}$$

The geometric mean decomposition is obtained by running the following code:

```
> tfpmb_oo_base_ccd = deatfpmb(X, Y, 'orient', 'oo', 'period', 'base', ...
' decomp', 'ccd', 'names', STATE_NAME)
```

If the base period technology exhibits CRS then the corresponding MBPI and its decomposition is obtained by setting the parameter decomp to crs:

```
> tfpmb_oo_base_complete = deatfpmb(X, Y, 'orient', 'oo', 'period', 'base', ...
' decomp', 'crs', 'names', STATE_NAME)
```

4.1.3. The comparison-period MBPI

Similar decompositions of the MBPI can be obtained if the comparison period technology is selected as benchmark,  $MB^1(x^1, y^1, x^0, y^0; \bar{x}, \bar{y})$ . The geometric mean of the decompositions along 'Paths C and D' is obtained by setting the period parameter to comparison. Related decompositions are obtained by setting the decomp parameter to ccd or crs.

```
> tfpmb_oo_comp_complete = deatfpmb(X, Y, 'orient', 'oo', 'period', 'comparison', ...
'decom', 'complete', 'names', STATE_NAME)
```

4.1.4. The geometric-mean MBPI

The base- and comparison-period MBPIs yield different results unless the benchmark technologies satisfy extremely restrictive conditions. As we have seen, there are four decompositions of the productivity index  $MB^t(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}})$  from an output orientation. Expressions (19) and (20) provide two decompositions of  $MB^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}})$ . There are two similar decompositions of  $MB^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}})$ . By taking their geometric mean, we obtain a decomposition of the geometric mean index  $MB(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}}) = [MB^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}})MB^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}})]^{1/2}$ . Thus,

$$\begin{aligned}
 MB(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}}) &= EC_0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times [TC_0^{0.1}(\mathbf{x}^1, \mathbf{y}^1) \times TC_0^{0.1}(\mathbf{x}^0, \mathbf{y}^0)]^{1/2} \times \\
 &\underbrace{\left[ \underbrace{\left( \frac{D_0^0(\mathbf{x}^0, \mathbf{y}^0)}{D_0^0(\mathbf{x}^1, \mathbf{y}^0)} \frac{D_i^0(\mathbf{x}^0, \bar{\mathbf{y}})}{D_i^0(\mathbf{x}^1, \bar{\mathbf{y}})} \right)}_{\text{Third term Path A. SEC}} \times \underbrace{\left( \frac{D_0^1(\mathbf{x}^0, \mathbf{y}^0)}{D_0^1(\mathbf{x}^1, \mathbf{y}^0)} \frac{D_i^1(\mathbf{x}^0, \bar{\mathbf{y}})}{D_i^1(\mathbf{x}^1, \bar{\mathbf{y}})} \right)}_{\text{Third term Path C. SEC}} \right]}_{\text{Third terms paths AC, SEC}}^{1/4} \times \\
 &\underbrace{\left[ \underbrace{\left( \frac{D_0^0(\mathbf{x}^0, \mathbf{y}^1)}{D_0^0(\mathbf{x}^1, \mathbf{y}^1)} \frac{D_i^0(\mathbf{x}^0, \bar{\mathbf{y}})}{D_i^0(\mathbf{x}^1, \bar{\mathbf{y}})} \right)}_{\text{Third term Path B. SEC}} \times \underbrace{\left( \frac{D_0^1(\mathbf{x}^0, \mathbf{y}^1)}{D_0^1(\mathbf{x}^1, \mathbf{y}^1)} \frac{D_i^1(\mathbf{x}^0, \bar{\mathbf{y}})}{D_i^1(\mathbf{x}^1, \bar{\mathbf{y}})} \right)}_{\text{Third term Path D. SEC}} \right]}_{\text{Third terms paths BD, SEC}}^{1/4} \times \\
 &\times \underbrace{\left[ \underbrace{\left( \frac{D_0^0(\bar{\mathbf{x}}, \mathbf{y}^1)}{D_0^0(\bar{\mathbf{x}}, \mathbf{y}^0)} \frac{D_0^0(\mathbf{x}^1, \mathbf{y}^0)}{D_0^0(\mathbf{x}^1, \mathbf{y}^1)} \right)}_{\text{Fourth term Path A. OME}} \times \underbrace{\left( \frac{D_0^1(\bar{\mathbf{x}}, \mathbf{y}^1)}{D_0^1(\bar{\mathbf{x}}, \mathbf{y}^0)} \frac{D_0^1(\mathbf{x}^1, \mathbf{y}^0)}{D_0^1(\mathbf{x}^1, \mathbf{y}^1)} \right)}_{\text{Fourth term Path C. OME}} \right]}_{\text{Fourth terms paths AC, OME}}^{1/4} \times \\
 &\underbrace{\left[ \underbrace{\left( \frac{D_0^0(\bar{\mathbf{x}}, \mathbf{y}^1)}{D_0^0(\bar{\mathbf{x}}, \mathbf{y}^0)} \frac{D_0^0(\mathbf{x}^0, \mathbf{y}^0)}{D_0^0(\mathbf{x}^0, \mathbf{y}^1)} \right)}_{\text{Fourth term Path B. OME}} \times \underbrace{\left( \frac{D_0^1(\bar{\mathbf{x}}, \mathbf{y}^1)}{D_0^1(\bar{\mathbf{x}}, \mathbf{y}^0)} \frac{D_0^1(\mathbf{x}^0, \mathbf{y}^0)}{D_0^1(\mathbf{x}^0, \mathbf{y}^1)} \right)}_{\text{Fourth term Path D. OME}} \right]}_{\text{Fourth terms paths BD, OME}}^{1/4} .
 \end{aligned} \tag{23}$$

The first row of expression (23) delivers the efficiency change and technological change effects, respectively. The second and third rows together correspond to the radial scale plus input mix effect (gathering the third factors in the component decompositions), and the fourth and fifth rows together measure the output mix effect (gathering the fourth factors in the component decompositions). This decomposition is obtained by setting in the deatfpmb(X, Y, ...) function the period parameter to geomean.

```
> tfpmb_oo_geo_complete = deatfpmb(X, Y, 'orient', 'oo', 'period', 'geomean', ...
'decom', 'complete', 'names', STATE_NAME);
> tfpdisp(tfpmb_oo_geo_complete)
```

-----  
Total Factor Productivity (TFP)

DMUs: 48  
Inputs: 4      Outputs: 3  
Model: tfp-mb  
Orientation: oo (Output orientated)  
Returns to scale: vrs (Variable)

Moorsteen--Bjurek index:  
Reference period is Geometric mean

DMU	MB	EC	TC	SEC	OME
AL	1.5915	1.3105	1.2565	0.9724	0.9939
AR	1.9443	1.3527	1.3553	1.0102	1.0498
AZ	1.3206	0.9722	1.5511	0.8833	0.9915
...					
WI	1.2247	1.0397	1.1133	0.9781	1.0819
WV	0.7524	1.0020	0.7787	0.9876	0.9764
WY	0.8427	0.8509	1.0042	0.9779	1.0084
Mean	1.4796	1.1101	1.3352	0.9762	1.0130
Std	0.4024	0.1823	0.3016	0.0444	0.0430
Max	2.7258	1.4695	2.0343	1.0701	1.1174
Min	0.4669	0.7703	0.6472	0.8545	0.9305

-----

The separate TC, SEC, and OME components can be recovered from the previous decompositions under the base or comparison specification. The related decompositions can be obtained by changing the argument complete in the decomp parameter to ccd or crs.

Again, for US agriculture, it is possible to compare productivity trends delivered by the geometric-mean-output-orientated Moorsteen-Bjurek index and its Malmquist counterpart. As with the base-period perspective comparison, both sets of results are highly correlated with a Spearman coefficient equal to  $\rho = 0.8763$ , statistically significant at the 1% level. Indeed, average productivity growth according to the Moorsteen-Bjurek index amounts 47.96%, slightly smaller than the Malmquist index at 54.97%. Also the maximum and minimum values are very similar, corresponding in both cases to Delaware (DE) and Oklahoma (OK), respectively. The only difference between the two indices is in the scale and output mix effects, but these differences are rather small.

4.2. The input-orientated decomposition of MBPI

The counterpart decomposition of the MBPI – see expression (19) – taking the technology of the base period as benchmark, along ‘Path E’ is given by

$$MB^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}}) = EC_i(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_i^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times \left( \frac{D_i^0(\mathbf{x}^1, \mathbf{y}^0)}{D_i^0(\mathbf{x}^0, \mathbf{y}^0)} \frac{D_i^0(\mathbf{x}^0, \bar{\mathbf{y}})}{D_i^0(\mathbf{x}^1, \bar{\mathbf{y}})} \right) \times \left( \frac{D_o^0(\bar{\mathbf{x}}, \mathbf{y}^1)}{D_o^0(\bar{\mathbf{x}}, \mathbf{y}^0)} \frac{D_i^0(\mathbf{x}^1, \mathbf{y}^1)}{D_i^0(\mathbf{x}^1, \mathbf{y}^0)} \right). \tag{24}$$

The factors on the right-hand side of the equality sign represent, respectively, technical efficiency change, technological change (conditional on  $(\mathbf{x}^1, \mathbf{y}^1)$ ), the input mix effect (conditional on  $\mathbf{y}^0$ ), and the radial scale plus output mix effect (conditional on  $\mathbf{x}^1$ ). The decomposition along ‘Path F’ is the same except that the input mix effect is conditional on  $\mathbf{y}^1$  and the radial scale plus output mix effect is conditional on  $\mathbf{x}^0$ . The geometric mean of the two decompositions is calculated setting the orient parameter to io, and the period parameter to base in the deatfpmb(X, Y, ...) function.

```
> tfpmb_io_base_complete = deatfpmb(X, Y, 'orient', 'io', 'period', 'base', ...
'decomp', 'complete', 'names', STATE_NAME)
```

Related decompositions are obtained by changing the decomp parameter to ccd or crs.

Alternatively, the input-orientated decompositions of the comparison-period MBPI along ‘Path G’ and ‘Path H’, and their geometric mean, are obtained by setting the period parameter to comparison. Changing the decomp parameter to ccd or crs yields the related decompositions.

```
> tfpmb_io_comp_complete = deatfpmb(X, Y, 'orient', 'io', 'period', 'comparison', ...
'decomp', 'complete', 'names', STATE_NAME)
```

All in all, there are four decompositions of the productivity index  $MB^t(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}})$  from the input orientation: two from the base period perspective, and two from the comparison period perspective. By taking the geometric mean of these four decompositions we obtain a decomposition of the geometric mean index  $MB(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}}) = [MB^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}})MB^1(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \bar{\mathbf{x}}, \bar{\mathbf{y}})]^{1/2}$ . This corresponds to the decomposition in expression (23), but is based on input distance functions instead of output distance functions.

The geometric mean decomposition is obtained by setting the period parameter to `geomean`.

```
> tfpmb_io_geo_complete = deatfpmb(X, Y, 'orient', 'io', 'period', 'geomean', ...
'decomp', 'complete', 'names', STATE_NAME)
```

## 5. Price-weighted productivity indices (Fisher)

The generic definition of a price-weighted productivity index comparing  $(\mathbf{x}^1, \mathbf{y}^1)$  to  $(\mathbf{x}^0, \mathbf{y}^0)$ , conditional on input prices  $\mathbf{w}^t$  and output prices  $\mathbf{p}^t$ , is

$$PROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{w}^t, \mathbf{p}^t) \equiv \frac{\mathbf{p}^t \cdot \mathbf{y}^1 / \mathbf{w}^t \cdot \mathbf{x}^1}{\mathbf{p}^t \cdot \mathbf{y}^0 / \mathbf{w}^t \cdot \mathbf{x}^0} = \frac{\mathbf{p}^t \cdot \mathbf{y}^1 / \mathbf{p}^t \cdot \mathbf{y}^0}{\mathbf{w}^t \cdot \mathbf{x}^1 / \mathbf{w}^t \cdot \mathbf{x}^0}. \quad (25)$$

Defined as a ratio of Lowe quantity indices of output and input, respectively, it can also be seen as a ratio of the productivity levels of period 1 and period 0, respectively. Period  $t$  serves as benchmark period, determining the prices with which quantities are multiplied. For  $t = 0$  we obtain the Laspeyres productivity index, and for  $t = 1$  the Paasche productivity index. The geometric mean of these two is the Fisher productivity index.

To calculate these indices, users must provide not only input and output quantities ( $X$  and  $Y$  matrices) but also corresponding prices ( $W$  and  $P$  matrices), which are three-dimensional arrays, with the third dimension corresponding to the time periods  $t = 0, 1, \dots, T$ . Note that in principle, the price-weighted indices could be calculated using any set of prices provided by the user. However, if prices are chosen as above, the toolbox calculates the Laspeyres, Paasche, and Fisher indices by default.

### 5.1. The output-orientated decomposition

#### 5.1.1. The Laspeyres PI

Following 'Path A', the decomposition of the Laspeyres PI is

$$PROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{w}^0, \mathbf{p}^0) = EC_0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_0^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times \left[ \frac{D_0^0(\mathbf{x}^0, \mathbf{y}^0) \mathbf{w}^0 \cdot \mathbf{x}^0}{D_0^0(\mathbf{x}^1, \mathbf{y}^0) \mathbf{w}^0 \cdot \mathbf{x}^1} \right] \times \left[ \frac{D_0^0(\mathbf{x}^1, \mathbf{y}^0) \mathbf{p}^0 \cdot \mathbf{y}^1}{D_0^0(\mathbf{x}^1, \mathbf{y}^1) \mathbf{p}^0 \cdot \mathbf{y}^0} \right]. \quad (26)$$

At the right-hand side, the first factor measures technical efficiency change, and the second factor measures technological change. [Balk and Zofio \(2018\)](#) show that the third factor corresponds to the scale (including input mix) effect, while the fourth factor corresponds to the output mix effect. Following the alternative 'Path B' delivers

$$PROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{w}^0, \mathbf{p}^0) = EC_0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_0^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times \left[ \frac{D_0^0(\mathbf{x}^0, \mathbf{y}^1) \mathbf{w}^0 \cdot \mathbf{x}^0}{D_0^0(\mathbf{x}^1, \mathbf{y}^1) \mathbf{w}^0 \cdot \mathbf{x}^1} \right] \times \left[ \frac{D_0^0(\mathbf{x}^0, \mathbf{y}^0) \mathbf{p}^0 \cdot \mathbf{y}^1}{D_0^0(\mathbf{x}^0, \mathbf{y}^1) \mathbf{p}^0 \cdot \mathbf{y}^0} \right], \quad (27)$$

which differs from expression (26) in the conditioning variables of the scale and output mix effects. Taking the geometric mean of these two expressions yields

$$PROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{w}^0, \mathbf{p}^0) = EC_0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_0^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times \left[ \left( \frac{D_0^0(\mathbf{x}^0, \mathbf{y}^0) D_0^0(\mathbf{x}^0, \mathbf{y}^1)}{D_0^0(\mathbf{x}^1, \mathbf{y}^0) D_0^0(\mathbf{x}^1, \mathbf{y}^1)} \right)^{1/2} \frac{\mathbf{w}^0 \cdot \mathbf{x}^0}{\mathbf{w}^0 \cdot \mathbf{x}^1} \right] \times \left[ \left( \frac{D_0^0(\mathbf{x}^1, \mathbf{y}^0) D_0^0(\mathbf{x}^0, \mathbf{y}^0)}{D_0^0(\mathbf{x}^1, \mathbf{y}^1) D_0^0(\mathbf{x}^0, \mathbf{y}^1)} \right)^{1/2} \frac{\mathbf{p}^0 \cdot \mathbf{y}^1}{\mathbf{p}^0 \cdot \mathbf{y}^0} \right]. \quad (28)$$

This decomposition is computed by the `deatfprod(X, Y, ...)` function. The `orient` parameter should be set to the output orientation `oo`, and the `period` parameter to `base`. The decomposition parameter `decomp` should be `complete`.

```
> tfprod_oo_base_complete = deatfprod(X, Y, W, P, 'orient', 'oo', 'period', 'base', ...
'decomp', 'complete', 'names', STATE_NAME);
> tfpdisp(tfprod_oo_base_complete)
```

```

-----
Total Factor Productivity (TFP)
DMUs: 48      Inputs: 4      Outputs: 3
Model: tfp-prod
Orientation: oo (Output orientated)
Returns to scale: vrs (Variable)
Price-weighted productivity index: Reference period is base period
-----

```

DMU	Prod	EC	TC	SEC	OME	SEC_100	OME_110	SEC_101	OME_010
AL	0.9652	1.3105	2.6698	0.2740	1.0068	0.2740	1.0068	0.2740	1.0068
AR	1.2647	1.3527	3.2174	0.3394	0.8561	0.3515	0.8268	0.3278	0.8865
AZ	1.0194	0.9722	3.5293	0.4859	0.6114	0.4806	0.6182	0.4913	0.6047
...									
WI	1.2489	1.0397	1.8528	0.5792	1.1194	0.5517	1.1752	0.6081	1.0662
WV	1.0936	1.0020	0.8737	1.0810	1.1555	1.1263	1.1090	1.0375	1.2040
WY	0.8745	0.8509	1.4799	0.6139	1.1311	0.6495	1.0693	0.5804	1.1965
Mean	1.2031	1.1101	2.9944	0.5478	0.7794	0.5494	0.7860	0.5482	0.7575
Std	0.2458	0.1823	1.2485	0.1798	0.2430	0.1950	0.2620	0.1698	0.2484
Max	1.7575	1.4695	6.6537	1.0810	1.3895	1.1263	1.6949	1.0375	1.2040
Min	0.6712	0.7703	0.8737	0.2583	0.3750	0.2433	0.3850	0.2592	0.1984

Again, the output successively shows the Laspeyres PI (PROD), EC, TC, the scale (plus input mix) factor SEC, and the output mix factor OME. The last four columns show the separate values of the scale and output mix effects along 'Path A' and 'Path B'. They are identified by the time superscripts of the input and output arguments; for instance, SEC\_100 identifies the third right-hand side factor in expression (26), which is the counterpart to  $SEC_0^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0)$  in expression (3). In particular cases distance functions may be infeasible, and therefore some of the factors in the above decompositions cannot be calculated. This is why the toolbox directly computes the Laspeyres PI by using expression (25).

For the US agriculture, we can compare the results of the standard Laspeyres PI with the previous quantities-only counterparts. Using prices to aggregate inputs and outputs results in productivity growth magnitudes which are much smaller. Over the whole period Laspeyres productivity (PROD) increased by 20.31% on average, which is a fraction (about ten percent) of the 215.99% and 229.52% increase of the base-period Malmquist and Moorsteen-Bjurek productivity indices. Since the EC and TC factors are the same, we see that the SEC and OME effects are responsible for the reduction in productivity. Indeed, these terms are responsible for a negative contribution to the tune of -45.55% and -22.06%, respectively, over forty-four years. It is then clear that the overall picture greatly changes when moving from quantities-only to price-based productivity indices, and different interpretations and policy conclusions could be drawn depending on the specific definitions. There are also some changes at the individual level. Oregon (OR) now leads the productivity growth with a 75.75% increase, and Oklahoma (OK) resides at the bottom with a 32.88% decrease. However, the rankings are relatively similar, exhibiting positive correlation. The Spearman correlation between the Laspeyres and Malmquist indices is  $\rho = 0.3088$ , while that between the Laspeyres and Moorsteen-Bjurek indices is  $\rho = 0.3067$ , both statistically significant at the usual 5% level.

5.1.2. Related decompositions of the Laspeyres PI

Using expression (6), expression (26) can be simplified to

$$PROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{w}^0, \mathbf{p}^0) = M_0^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times \left[ \frac{D_0^0(\mathbf{x}^0, \mathbf{y}^0) \mathbf{w}^0 \cdot \mathbf{x}^0}{D_0^0(\mathbf{x}^1, \mathbf{y}^0) \mathbf{w}^0 \cdot \mathbf{x}^1} \right] \times \left[ \frac{D_0^0(\mathbf{x}^1, \mathbf{y}^0) \mathbf{p}^0 \cdot \mathbf{y}^1}{D_0^0(\mathbf{x}^1, \mathbf{y}^1) \mathbf{p}^0 \cdot \mathbf{y}^0} \right]. \tag{29}$$

The same holds for expressions (27) and (28). To obtain such a decomposition, the decomp parameter must be set to ccd, combined with the base, comparison, or geometric mean perspective. For the Laspeyres PI we would run

```

> tfprod_oo_base_ccd = deatfprod(X, Y, W, P, 'orient', 'oo', 'period', 'base', ...
'decom', 'ccd', 'names', STATE_NAME)

```

If the base period technology is characterized by CRS, then the corresponding decomposition is obtained by setting the parameter decomp to crs.

```

> tfprod_oo_base_crs = deatfprod(X, Y, W, P, 'orient', 'oo', 'period', 'base', ...
'decom', 'crs', 'names', STATE_NAME)

```

5.1.3. The Paasche PI

Following "Path C" and 'Path D' and averaging the resulting decompositions, the decomposition of the Paasche PI (that is, expression (25) with  $t = 1$ ) reads

$$PROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{w}^1, \mathbf{p}^1) = EC_0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_0^{1,0}(\mathbf{x}^0, \mathbf{y}^0) \times \left[ \left( \frac{D_0^1(\mathbf{x}^0, \mathbf{y}^0) D_0^1(\mathbf{x}^0, \mathbf{y}^1)}{D_0^1(\mathbf{x}^1, \mathbf{y}^0) D_0^1(\mathbf{x}^1, \mathbf{y}^1)} \right)^{1/2} \frac{\mathbf{w}^1 \cdot \mathbf{x}^0}{\mathbf{w}^1 \cdot \mathbf{x}^1} \right] \times \left[ \left( \frac{D_0^1(\mathbf{x}^1, \mathbf{y}^0) D_0^1(\mathbf{x}^0, \mathbf{y}^0)}{D_0^1(\mathbf{x}^1, \mathbf{y}^1) D_0^1(\mathbf{x}^0, \mathbf{y}^1)} \right)^{1/2} \frac{\mathbf{p}^1 \cdot \mathbf{y}^1}{\mathbf{p}^1 \cdot \mathbf{y}^0} \right], \tag{30}$$

To obtain this decomposition the period parameter must be set to comparison in the `deatfprod(X, Y, ...)` function. For specific DMUs some factors may be infeasible to compute. The toolbox, however, computes the Paasche PI directly rather than as a product of the components.

```
> tfprod_oo_comp_complete = deatfprod(X, Y, W, P, 'orient', 'oo', 'period', 'comparison', ...
'decomp', 'complete', 'names', STATE_NAME)
```

The first two factors at the right-hand side of the equality sign in expression (30) correspond to the CCD index  $M_0^1(x^1, y^1, x^0, y^0)$ , defined by expression (14). The related decomposition is obtained by changing the `decomp` parameter from `complete` to `ccd`. Also, if CRS holds for the comparison period technology, then the decomposition is obtained by setting the `decomp` parameter to `crs`.

5.1.4. The Fisher PI

The Fisher PI is the geometric mean of the Laspeyres and Paasche PIs. By merging expressions (28) and (30) its decomposition is

$$PROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{w}^1, \mathbf{p}^1, \mathbf{w}^0, \mathbf{p}^0) \equiv \frac{Q^F(\mathbf{p}^1, \mathbf{y}^1, \mathbf{p}^0, \mathbf{y}^0)}{Q^F(\mathbf{w}^1, \mathbf{x}^1, \mathbf{w}^0, \mathbf{x}^0)} =$$

$$EC_0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times [TC_0^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times TC_0^{0,1}(\mathbf{x}^0, \mathbf{y}^0)]^{1/2} \times$$

$$\left[ \left( \frac{D_0^0(\mathbf{x}^0, \mathbf{y}^0) D_0^0(\mathbf{x}^0, \mathbf{y}^1) D_0^1(\mathbf{x}^0, \mathbf{y}^0) D_0^1(\mathbf{x}^0, \mathbf{y}^1)}{D_0^0(\mathbf{x}^1, \mathbf{y}^0) D_0^0(\mathbf{x}^1, \mathbf{y}^1) D_0^1(\mathbf{x}^1, \mathbf{y}^0) D_0^1(\mathbf{x}^1, \mathbf{y}^1)} \right)^{1/4} \frac{1}{Q^F(\mathbf{w}^1, \mathbf{x}^1, \mathbf{w}^0, \mathbf{x}^0)} \right] \times$$

$$\left[ \left( \frac{D_0^0(\mathbf{x}^1, \mathbf{y}^0) D_0^0(\mathbf{x}^0, \mathbf{y}^0) D_0^1(\mathbf{x}^1, \mathbf{y}^0) D_0^1(\mathbf{x}^0, \mathbf{y}^0)}{D_0^0(\mathbf{x}^1, \mathbf{y}^1) D_0^0(\mathbf{x}^0, \mathbf{y}^1) D_0^1(\mathbf{x}^1, \mathbf{y}^1) D_0^1(\mathbf{x}^0, \mathbf{y}^1)} \right)^{1/4} Q^F(\mathbf{p}^1, \mathbf{y}^1, \mathbf{p}^0, \mathbf{y}^0) \right], \tag{31}$$

where  $Q^F(\mathbf{p}^1, \mathbf{y}^1, \mathbf{p}^0, \mathbf{y}^0) \equiv [(\mathbf{p}^0 \cdot \mathbf{y}^1 / \mathbf{p}^0 \cdot \mathbf{y}^0)(\mathbf{p}^1 \cdot \mathbf{y}^1 / \mathbf{p}^1 \cdot \mathbf{y}^0)]^{1/2}$  is the Fisher output quantity index, and  $Q^F(\mathbf{w}^1, \mathbf{x}^1, \mathbf{w}^0, \mathbf{x}^0) \equiv [(\mathbf{w}^0 \cdot \mathbf{x}^1 / \mathbf{w}^0 \cdot \mathbf{x}^0)(\mathbf{w}^1 \cdot \mathbf{x}^1 / \mathbf{w}^1 \cdot \mathbf{x}^0)]^{1/2}$  the Fisher input quantity index.

This decomposition is obtained by calling the `deatfprod(X, Y, ...)` function with the period parameter set to `geomean`, the orientation parameter set to `oo`, and the `decomp` parameter to `complete`.

```
> tfprod_oo_geo_complete = deatfprod(X, Y, W, P, 'orient', 'oo', 'period', 'geomean', ...
'decomp', 'complete', 'names', STATE_NAME);
> tfpdisp(tfprod_oo_geo_complete)
```

```
-----
Total Factor Productivity (TFP)
DMUs: 48      Inputs: 4      Outputs: 3
Model: tfp-prod
Orientation: oo (Output orientated)
Returns to scale: vrs (Variable)
Price-weighted productivity index: Reference period is Geometric mean
-----
```

DMU	Prod	EC	TC	SEC	OME
AL	1.3021	1.3105	1.2565	0.7849	1.0074
AR	1.6309	1.3527	1.3553	0.9269	0.9597
AZ	1.1197	0.9722	1.5511	0.8570	0.8664
...					
WI	1.5886	1.0397	1.1133	1.1604	1.1828
WV	1.2819	1.0020	0.7787	1.5641	1.0504
WY	0.9053	0.8509	1.0042	0.9775	1.0838
-----					
Mean	1.3951	1.1101	1.3352	1.0078	0.9556
Std	0.2888	0.1823	0.3016	0.1740	0.1475
Max	2.0273	1.4695	2.0343	1.5641	1.5567
Min	0.7415	0.7703	0.6472	0.6579	0.6081
-----					

The output of the function successively shows the Fisher PI (PROD), EC, the geometric mean of the TC factors, the geometric mean of the four separate SEC effects as presented in the third row of expression (31), and the geometric mean of the four separate OME effects as presented in the fourth row of the same expression. Each of these four separate terms in SEC or OME corresponds to a specific path, A, B, C or D, which can be recovered by running the base and comparison period decompositions as explained above. To ensure that the Fisher index is reported, regardless of the fact that some distance function values may be infeasible to compute, the toolbox calculates the geometric mean of the Laspeyres and Paasche PIs.

It is interesting to compare the results obtained for the Fisher index to those of the geometric-mean Malmquist and Moorsteen-Bjurek productivity indices (MPI and MBPI). While the differences between the last two indices were minimal, with high correlation coefficients,



this is not the case when compared to the Fisher index. However, the values in this case are in the same order of magnitude, as opposed to the base period counterparts where the spread between the indices was rather large. Average US agricultural productivity change measured by the Fisher index is 39.51%, lower than the MPI and MBPI indices, 54.94% and 47.96%, respectively. The range of productivity change is also rather similar, with the best and worst performing states being Rhode Island (RI) and Oklahoma (OK), the latter state consistently ranking last as in the previous two cases. The dissimilarity across rankings for the geometric-mean indices is captured by the Spearman coefficients:  $\rho = 0.5199$  for Fisher-MPI, and  $\rho = 0.5680$  for Fisher-MBPI, both statistically significant at the 1% confidence level. These values are notably smaller than the coefficient for MPI-MBPI,  $\rho = 0.8763$ . Notice that, as in the base-period case, SEC and OME are responsible for the differences between the indices.

Finally, expression (31) with the EC and TC factors replaced by the geometric-mean CCD index is obtained by changing the decomp parameter to `ccd`. Under CRS the decomposition is obtained by setting the parameter to `crs`.

### 5.2. The input-orientated decomposition

The Laspeyres, Paasche, and Fisher productivity indices can also be decomposed from the input-orientated perspective.

Along 'Path E' and 'Path F' one obtains decompositions for the Laspeyres PI similar to expressions (26) and (27), except that input distance functions are used instead of output distance functions. Their geometric mean, corresponding to expression (28), is

$$PROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{w}^0, \mathbf{p}^0) = EC_i(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_i^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times \left[ \left( \frac{D_i^0(\mathbf{x}^1, \mathbf{y}^0) D_i^0(\mathbf{x}^1, \mathbf{y}^1)}{D_i^0(\mathbf{x}^0, \mathbf{y}^0) D_i^0(\mathbf{x}^0, \mathbf{y}^1)} \right)^{1/2} \frac{\mathbf{w}^0 \cdot \mathbf{x}^0}{\mathbf{w}^0 \cdot \mathbf{x}^1} \right] \times \left[ \left( \frac{D_i^0(\mathbf{x}^1, \mathbf{y}^1) D_i^0(\mathbf{x}^0, \mathbf{y}^1)}{D_i^0(\mathbf{x}^1, \mathbf{y}^0) D_i^0(\mathbf{x}^0, \mathbf{y}^0)} \right)^{1/2} \frac{\mathbf{p}^0 \cdot \mathbf{y}^1}{\mathbf{p}^0 \cdot \mathbf{y}^0} \right]. \tag{32}$$

The first two factors capture technical efficiency change and technological change, respectively. The third and fourth factors measure the input-mix effect and the scale (including output-mix) effect, respectively.

This decomposition is obtained through the `deatfprod(X, Y, ...)` function by setting the `orient` parameter to the input orientation `io`, setting the period parameter to `base`, and setting the decomposition parameter `decomp` to `complete`.

```
> tfprod_io_base_complete = deatfprod(X, Y, W, P, 'orient', 'io', 'period', 'base', ...
'decom', 'complete', 'names', STATE_NAME)
```

To see the CCD index  $M_i^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)$  as part of the decomposition, the `decomp` parameter must be set to `ccd`. If CRS holds for the base period technology (determined after running the function `deateststrtsm(X, Y, ...)` with the `orient` parameter set to `io`), then setting the `decomp` parameter to `crs` provides the correct decomposition. In this case, the third factor measures the input-mix effect, but the fourth factor measures only the output-mix effect, since under CRS the scale effect vanishes.

By combining the results of following 'Path G' and 'Path H' one obtains the input-orientated decomposition of the Paasche PI. In the `deatfprod(X, Y, ...)` function the period parameter must then be set to `comparison`. The related decompositions require setting the `decomp` parameter to `ccd` or `crs`.

```
> tfprod_io_comp_complete = deatfprod(X, Y, W, P, 'orient', 'io', 'period', 'comparison', ...
'decom', 'complete', 'names', STATE_NAME)
```

Finally, by taking the geometric mean of the decomposition in expression (32) and its comparison period counterpart, one obtains the input-distance-function based decomposition of the Fisher PI. This is calculated by setting the period parameter set to `geomean`, the orientation parameter to `io`, and the `decomp` parameter to `complete`.

```
> tfprod_io_geo_complete = deatfprod(X, Y, W, P, 'orient', 'io', 'period', 'geomean', ...
'decom', 'complete', 'names', STATE_NAME)
```

The decomposition featuring the geometric-mean CCD index is obtained by setting the `decomp` parameter to `ccd`. The CRS specification, valid if both technologies indeed exhibit CRS, is obtained by setting this parameter to `crs`.

## 6. Share-weighted productivity indices (Törnqvist)

A share-weighted productivity index is defined as a weighted geometric mean of output quantity changes divided by a weighted geometric mean of input quantity changes. In particular, individual input quantity changes,  $x_n^1/x_n^0$ , and output quantity changes,  $y_m^1/y_m^0$ , are weighted by their respective cost and revenue shares in a certain period  $t$ . The generic definition is

$$GPROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{s}^t, \mathbf{u}^t) \equiv \frac{\prod_{m=1}^M (y_m^1/y_m^0)^{u_m^t}}{\prod_{n=1}^N (x_n^1/x_n^0)^{s_n^t}}, \tag{33}$$

where  $s_n^t \equiv w_n^t x_n^t / w^t \cdot \mathbf{x}^t > 0$  ( $n = 1, \dots, N$ ),  $\sum_{n=1}^N s_n^t = 1$ ,  $u_m^t \equiv p_m^t y_m^t / p^t \cdot \mathbf{y}^t > 0$  ( $m = 1, \dots, M$ ),  $\sum_{m=1}^M u_m^t = 1$ . For  $t = 0$  the index is called Geolaspeyres, and for  $t = 1$  the index is called GeoPaasche. The geometric mean of the two indices is the Törnqvist productivity index. This index can also be obtained by considering the 'average' viewpoint for the cost and revenue shares,  $\mathbf{s}^{0,1} \equiv (\mathbf{s}^0 + \mathbf{s}^1)/2$  and  $\mathbf{u}^{0,1} \equiv (\mathbf{u}^0 + \mathbf{u}^1)/2$ .

6.1. The output-orientated decomposition

6.1.1. The Geolaspeyres PI

Following 'Path A', the Geolaspeyres PI is decomposed as

$$GPROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{s}^0, \mathbf{u}^0) = EC_0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_0^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times \left[ \frac{D_o^0(\mathbf{x}^0, \mathbf{y}^0)}{D_o^0(\mathbf{x}^1, \mathbf{y}^1)} \prod_{n=1}^N (x_n^0/x_n^1)^{s_n^0} \right] \times \left[ \frac{D_o^0(\mathbf{x}^1, \mathbf{y}^0)}{D_o^0(\mathbf{x}^1, \mathbf{y}^1)} \prod_{m=1}^M (y_m^1/y_m^0)^{u_m^0} \right]. \tag{34}$$

The first two factors after the equality sign are the same as the first two factors of the Laspeyres PI in expression (26); that is, technical efficiency change and technological change, respectively. The third factor captures the scale (including input-mix) effect, and the fourth factor measures the output-mix effect.

Following the alternative 'Path B' delivers

$$GPROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{s}^0, \mathbf{u}^0) = EC_0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_0^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times \left[ \frac{D_o^0(\mathbf{x}^0, \mathbf{y}^1)}{D_o^0(\mathbf{x}^1, \mathbf{y}^1)} \prod_{n=1}^N (x_n^0/x_n^1)^{s_n^0} \right] \times \left[ \frac{D_o^0(\mathbf{x}^0, \mathbf{y}^0)}{D_o^0(\mathbf{x}^0, \mathbf{y}^1)} \prod_{m=1}^M (y_m^1/y_m^0)^{u_m^0} \right], \tag{35}$$

which differs from expression (34) in the scale and the output-mix effect. Taking the geometric average of these two expressions yields

$$GPROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{s}^0, \mathbf{u}^0) = EC_0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_0^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times \left[ \left( \frac{D_o^0(\mathbf{x}^0, \mathbf{y}^0) D_o^0(\mathbf{x}^0, \mathbf{y}^1)}{D_o^0(\mathbf{x}^1, \mathbf{y}^0) D_o^0(\mathbf{x}^1, \mathbf{y}^1)} \right)^{1/2} \prod_{n=1}^N (x_n^0/x_n^1)^{s_n^0} \right] \times \left[ \left( \frac{D_o^0(\mathbf{x}^1, \mathbf{y}^0) D_o^0(\mathbf{x}^0, \mathbf{y}^0)}{D_o^0(\mathbf{x}^1, \mathbf{y}^1) D_o^0(\mathbf{x}^0, \mathbf{y}^1)} \right)^{1/2} \prod_{m=1}^M (y_m^1/y_m^0)^{u_m^0} \right]. \tag{36}$$

The function to calculate these decompositions is `deatfgprod(X, Y, W, P, 'orient', 'oo', 'period', 'base', ...)`, with the `orient` parameter set to output orientation `oo` and the `period` parameter to base. The decomposition parameter `decomp` must be set to `complete`.

```
> tfpgprod_oo_base_complete = deatfgprod(X, Y, W, P, 'orient', 'oo', 'period', 'base', ...
'decomp', 'complete', 'names', STATE_NAME);
> tfpdisp(tfpgprod_oo_base_complete)
```

```
-----
Total Factor Productivity (TFP)
DMUs: 48      Inputs: 4      Outputs: 3
Model: tfp-gprod
Orientation: oo (Output orientated)
Returns to scale: vrs (Variable)
Share-weighted geometric productivity index: Reference period is base period
-----
```

DMU	GProd	EC	TC	SEC	OME	SEC_100	OME_110	SEC_101	OME_010
AL	1.2707	1.3105	2.6698	0.3631	1.0001	0.3631	1.0001	0.3631	1.0001
AR	1.7608	1.3527	3.2174	0.4858	0.8328	0.5031	0.8043	0.4692	0.8624
AZ	0.9867	0.9722	3.5293	0.5229	0.5500	0.5171	0.5561	0.5287	0.5440
...									
WI	1.4101	1.0397	1.8528	0.7004	1.0452	0.6671	1.0973	0.7353	0.9956
WV	1.2408	1.0020	0.8737	1.2657	1.1197	1.3188	1.0747	1.2148	1.1667
WY	0.9310	0.8509	1.4799	0.6807	1.0861	0.7201	1.0267	0.6435	1.1489
Mean	1.2425	1.1101	2.9944	0.6531	0.6996	0.6542	0.7052	0.6545	0.6786
Std	0.3042	0.1823	1.2485	0.1823	0.2924	0.2003	0.3080	0.1718	0.2928
Max	1.7866	1.4695	6.6537	1.2657	1.3294	1.3188	1.6216	1.2148	1.1667
Min	0.6293	0.7703	0.8737	0.3631	0.2263	0.3631	0.2370	0.3631	0.1104

```
-----
```

The output of the function successively shows the Geolaspeyres productivity index (GPROD), EC, TC, and the geometric means and component values of the SEC and OME effects corresponding to 'Path A' and 'Path B'. The components are identified by the time superscripts of the input and output arguments in each factor; for instance, SEC\_100 represents the third factor in expression (34), counterpart to  $SEC_0^0(\mathbf{x}^1, \mathbf{x}^0, \mathbf{y}^0)$  in expression (2). As calculation of some distance function values may be infeasible, the toolbox calculates the Geolaspeyres PI directly from the data.

The share-weighted results for US agriculture are similar to their price-weighted counterparts. For instance, the Geolaspeyres and the Laspeyres deliver on average 24.25% and 20.31%, respectively. Also the maximum and minimum values are very similar, but on this occasion corresponding to Georgia (GA) and Florida (FL). This means that the individual rankings differ substantially between the two indices. Thus, using prices or value shares to aggregate outputs and inputs makes a difference. Indeed, the Spearman correlation between the two is  $\rho = 0.5404$ , the lowest across all possible pairs so far, but still significant at the 5% level. However, the correlations between the Geolaspeyres and the Malmquist (MPI) or the Moorsteen-Bjurek (MBPI) indices are even lower:  $\rho = 0.078$  and  $\rho = 0.2224$ , respectively, and not statistically significant at the standard confidence levels. Since the technical efficiency change (EC) and technological change (TC) factors are the same, this shows the importance of the scale and output-mix effects. All this calls for caution on the part of practitioners when choosing a specific definition of productivity change. If prices are available and reliable, then using a popular definition may yield results which differ substantially from analogous quantities-only definitions.

6.1.2. Related decompositions of the GeoLaspeyres PI

Recalling expression (6), expression (34) can be written as

$$GPROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{s}^0, \mathbf{u}^0) = M_o^0(x^1, y^1, x^0, y^0) \times \left[ \frac{D_o^0(\mathbf{x}^0, \mathbf{y}^0)}{D_o^0(\mathbf{x}^1, \mathbf{y}^0)} \prod_{n=1}^N (x_n^0/x_n^1)^{s_n^0} \right] \times \left[ \frac{D_o^0(\mathbf{x}^1, \mathbf{y}^0)}{D_o^0(\mathbf{x}^1, \mathbf{y}^1)} \prod_{m=1}^M (y_m^1/y_m^0)^{u_m^0} \right]. \tag{37}$$

Similar simplifications hold for expressions (35) and (36). To obtain these decompositions, the decomp parameter must be set to ccd.

```
> tfpgprod_oo_base_ccd = deatfpgprod(X, Y, W, P, 'orient', 'oo', ...
'period', 'base', 'decomp', 'ccd', 'names', STATE_NAME))
```

If the base period technology exhibits CRS then one must set the parameter decomp to crs.

```
> tfpgprod_oo_base_crs = deatfpgprod(X, Y, W, P, 'orient', 'oo', 'period', 'base', ...
'decomp', 'crs', 'names', STATE_NAME);
```

6.1.3. The GeoPaasche PI

Along 'Path C' and 'Path D' we obtain decompositions of the GeoPaasche PI. Their geometric mean, the analogue to expression (36), reads

$$GPROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{s}^1, \mathbf{u}^1) = EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_o^{1,0}(\mathbf{x}^0, \mathbf{y}^0) \times \left[ \left( \frac{D_o^1(\mathbf{x}^0, \mathbf{y}^0) D_o^1(\mathbf{x}^0, \mathbf{y}^1)}{D_o^1(\mathbf{x}^1, \mathbf{y}^0) D_o^1(\mathbf{x}^1, \mathbf{y}^1)} \right)^{1/2} \prod_{n=1}^N (x_n^0/x_n^1)^{s_n^1} \right] \times \left[ \left( \frac{D_o^1(\mathbf{x}^1, \mathbf{y}^0) D_o^1(\mathbf{x}^0, \mathbf{y}^0)}{D_o^1(\mathbf{x}^1, \mathbf{y}^1) D_o^1(\mathbf{x}^0, \mathbf{y}^1)} \right)^{1/2} \prod_{m=1}^M (y_m^1/y_m^0)^{u_m^1} \right]. \tag{38}$$

To obtain this decompositions, in the deatfpgprod(X, Y, ...) function the period parameter must be set to comparison.

```
> tfpgprod_oo_comp_complete = deatfpgprod(X, Y, W, P, 'orient', 'oo', 'period', 'comparison', ...
'decomp', 'complete', 'names', STATE_NAME);
```

The first two factors can be aggregated to the CCD index  $M_o^1(x^1, y^1, x^0, y^0)$ , as defined by expression (14). The corresponding table can be obtained by changing the decomp parameter from complete to ccd. If CRS holds for the comparison period technology then results are obtained by choosing crs.

6.1.4. The Törnqvist PI

Merging expressions (36) and (38) yields a decomposition of the Törnqvist PI, the share-weighted analogue of the Fisher PI,

$$GPROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; (\mathbf{s}^0 + \mathbf{s}^1)/2, (\mathbf{u}^0 + \mathbf{u}^1)/2) \equiv \frac{Q^T(\mathbf{p}^1, \mathbf{y}^1, \mathbf{p}^0, \mathbf{y}^0)}{Q^T(\mathbf{w}^1, \mathbf{x}^1, \mathbf{w}^0, \mathbf{x}^0)} = EC_o(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times [TC_o^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times TC_o^{1,0}(\mathbf{x}^0, \mathbf{y}^0)]^{1/2} \times \left[ \left( \frac{D_o^0(\mathbf{x}^0, \mathbf{y}^0) D_o^0(\mathbf{x}^0, \mathbf{y}^1) D_o^1(\mathbf{x}^0, \mathbf{y}^0) D_o^1(\mathbf{x}^0, \mathbf{y}^1)}{D_o^0(\mathbf{x}^1, \mathbf{y}^0) D_o^0(\mathbf{x}^1, \mathbf{y}^1) D_o^1(\mathbf{x}^1, \mathbf{y}^0) D_o^1(\mathbf{x}^1, \mathbf{y}^1)} \right)^{1/4} \frac{1}{Q^T(\mathbf{w}^1, \mathbf{x}^1, \mathbf{w}^0, \mathbf{x}^0)} \right] \times \left[ \left( \frac{D_o^0(\mathbf{x}^1, \mathbf{y}^0) D_o^0(\mathbf{x}^0, \mathbf{y}^0) D_o^1(\mathbf{x}^1, \mathbf{y}^0) D_o^1(\mathbf{x}^0, \mathbf{y}^0)}{D_o^0(\mathbf{x}^1, \mathbf{y}^1) D_o^0(\mathbf{x}^0, \mathbf{y}^1) D_o^1(\mathbf{x}^1, \mathbf{y}^1) D_o^1(\mathbf{x}^0, \mathbf{y}^1)} \right)^{1/4} Q^T(\mathbf{p}^1, \mathbf{y}^1, \mathbf{p}^0, \mathbf{y}^0) \right], \tag{39}$$

where the Törnqvist output and input quantity indices are defined by  $Q^T(p^1, y^1, p^0, y^0) \equiv \prod_{m=1}^M (y_m^1/y_m^0)^{(u_m^0+u_m^1)/2}$  and  $Q^T(w^1, x^1, w^0, x^0) \equiv \prod_{n=1}^N (x_n^1/x_n^0)^{(s_n^0+s_n^1)/2}$ , respectively. To calculate this decomposition, in the deatfpgprod(X, Y, ...) function the user sets period to geomean, orientation to oo, and decomp to complete.

```
> tfpgprod_oo_geo_complete = deatfpgprod(X, Y, W, P, 'orient', 'oo', 'period', 'geomean', ...
'decomp', 'complete', 'names', STATE_NAME);
> tfpdisp(tfpgprod_oo_geo_complete)
```

```

-----
Total Factor Productivity (TFP)
DMUs: 48      Inputs: 4      Outputs: 3
Model: tfp-gprod
Orientation: oo (Output orientated)
Returns to scale: vrs (Variable)
Share-weighted geometric productivity index: Reference period is Geometric mean
-----

```

DMU	GProd	EC	TC	SEC	OME
AL	1.2503	1.3105	1.2565	0.7524	1.0091
AR	1.5819	1.3527	1.3553	0.8994	0.9594
AZ	1.0914	0.9722	1.5511	0.8351	0.8666
...					
WI	1.5470	1.0397	1.1133	1.1332	1.1795
WV	1.2946	1.0020	0.7787	1.5797	1.0503
WY	0.9071	0.8509	1.0042	0.9836	1.0793
-----					
Mean	1.3768	1.1101	1.3352	1.0005	0.9527
Std	0.2834	0.1823	0.3016	0.1812	0.1497
Max	1.9378	1.4695	2.0343	1.5797	1.5509
Min	0.7362	0.7703	0.6472	0.6457	0.5999
-----					

The output of the function successively shows the Törnqvist productivity index (GPROD), the technical efficiency change factor (EC), the geometric mean of the two technological change factors (TC), the geometric mean of the four component scale efficiency effects in the third row of expression (39) (SEC), and the geometric mean of the four component output-mix effects in the fourth row of the same expression (OME). Each of these four components corresponds to a specific path A, B, C, or D, and can be recovered by running the previous base and comparison period decompositions. The toolbox calculates the Törnqvist index as the geometric mean of the GeoLaspeyres and GeoPaasche indices. Some factors might be missing because of infeasible distance function values.

The three-factor decomposition exhibiting the geometric-mean CCD index is obtained by changing the `decomp` parameter to `ccd`. If technologies of base and comparison periods exhibit CRS, then the decomposition is calculated by setting the same parameter to `crs`.

We can now compare the results of the Törnqvist productivity index to those of the geometric-mean output-orientated Malmquist (MPI), Moorsteen-Bjurek (MBPI), and Fisher indices. Average productivity growth in US agriculture, 37.68%, is the lowest compared to the previous three: 54.94% (MPI), 47.96% (MBPI), and 39.51% (Fisher). All these indices share the same efficiency change (EC) and technological change (TC) factors, and therefore the differences come from the scale and output-mix effects (SEC) and (OME). In this case the states exhibiting the maximum and minimum values coincide with those observed for the Fisher index, that is, Rhode Island (RI) and Oklahoma (OK). As for the rankings, the Spearman correlation coefficient between the two is almost one,  $\rho = 0.9948$ , and statistically significant. Those for Törnqvist-MPI and Törnqvist-MBPI are lower:  $\rho = 0.4822$  and  $\rho = 0.5397$ , respectively.

## 6.2. The input-orientated decomposition

All the share-weighted productivity indices discussed in the previous subsection can also be decomposed using input distance functions. Along 'Path E' and 'Path F' one obtains expressions analogous to (34) and (35), but now using input distance functions. The geometric mean counterpart to expression (36) is

$$GPROD(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0; \mathbf{s}^0, \mathbf{u}^0) = EC_i(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0) \times TC_i^{1,0}(\mathbf{x}^1, \mathbf{y}^1) \times$$

$$\left[ \left( \frac{D_i^0(\mathbf{x}^1, \mathbf{y}^0) D_i^0(\mathbf{x}^1, \mathbf{y}^1)}{D_i^0(\mathbf{x}^0, \mathbf{y}^0) D_i^0(\mathbf{x}^0, \mathbf{y}^1)} \right)^{1/2} \prod_{n=1}^N (x_n^0/x_n^1)^{s_n^0} \right] \times \left[ \left( \frac{D_i^0(\mathbf{x}^1, \mathbf{y}^1) D_i^0(\mathbf{x}^0, \mathbf{y}^1)}{D_i^0(\mathbf{x}^1, \mathbf{y}^0) D_i^0(\mathbf{x}^0, \mathbf{y}^0)} \right)^{1/2} \prod_{m=1}^M (y_m^1/y_m^0)^{u_m^0} \right]. \quad (40)$$

The first two factors measure technical efficiency change and technological change, respectively. The third factor measures the input-mix effect and the fourth factor measures the scale (including output-mix) effect. In the `deatfgprod(X, Y, ...)` function, this decomposition requires setting the `orient` parameter to `io`, the `period` parameter to `base`, and the `decomp` parameter to `complete`.

```

> tfpgprod_io_base_complete = deatfgprod(X, Y, W, P, 'orient', 'io', 'period', 'base', ...
' decomp', 'complete', 'names', STATE_NAME)

```

To recover the CCD index  $M_i^0(\mathbf{x}^1, \mathbf{y}^1, \mathbf{x}^0, \mathbf{y}^0)$ , the `decomp` parameter must be set to `ccd`. If the base period technology exhibits CRS then setting the `decomp` parameter to `crs` calculates the appropriate decomposition.

Alternatively, the comparison-period analogue of expression (40), corresponding to the geometric mean of 'Path G' and 'Path H', is computed by setting the `period` parameter to `comparison`. The decompositions featuring the comparison-period CCD index, and the decomposition under CRS (of the comparison period technology), require setting the `decomp` parameter to `ccd` or `crs`, respectively.

```

> tfpgprod_io_comp_complete = deatfgprod(X, Y, W, P, 'orient', 'io', 'period', 'comparison', ...
' decomp', 'complete', 'names', STATE_NAME)

```

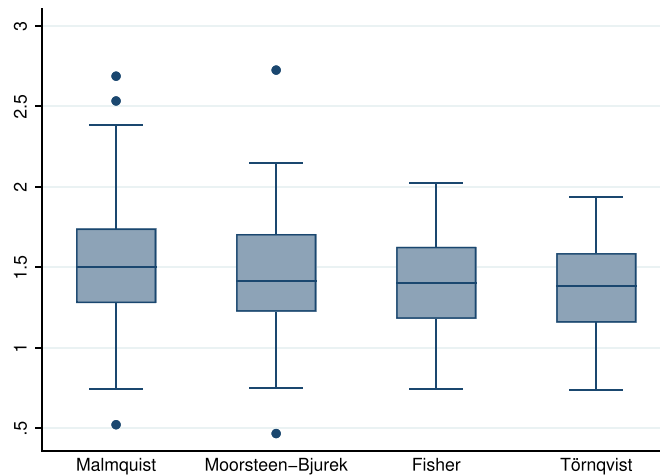


Fig. 2. Box-plots of the productivity distributions, geometric means.

Table 2

Correlation and distribution tests between TFP growth according to USDA and this paper.

	Pearson's $r$	Spearman's $\rho$	Kendall's $\tau$	$t$ test	Wilcoxon
Malmquist	0.5516*	0.6081*	0.4175*	10.2777*	5.703*
Moorsteen-Bjurek	0.6431*	0.7099*	0.5127*	12.6546*	5.567*
Fisher	0.8489*	0.8254*	0.6368*	24.3226*	6.031*
Törnqvist	0.8424*	0.8206*	0.6301*	24.3960*	6.031*

Note: \* denotes significance at the 0.001 level.

Finally, by taking the geometric mean of expression (40) and its comparison-period analogue, one obtains the input-orientated decomposition of the Törnqvist productivity index. This decomposition is obtained from the `deatfprod(X, Y, ...)` function by setting the period parameter to `geomean`, the orientation parameter to `io`, and the `decomp` parameter to `complete`.

```
> tfpgprod_io_geo_complete = deatfpgprod(X, Y, W, P, 'orient', 'io', 'period', 'geomean', ...
' decomp', 'complete', 'names', STATE_NAME)
```

The three-factor decomposition featuring the geometric-mean CCD indices, is obtained by setting the `decomp` parameter to `ccd`. The decomposition under CRS is obtained by setting this parameter to `crs`.

## 7. Empirical results: TFP growth in US agriculture

As for the empirical applications, the literature provides us with a number of implementations. However, most of these implementations, sometimes on microdata, sometimes on macrodata, appear to be partial, usually concerning only one or two decompositions at a time. The unique feature of this paper is that all the theoretically possible decompositions are applied to the same US agricultural dataset with information at the state level. In the particular case of our dataset, we are able to judge the extent to which the various measures and decompositions are empirically different. Focusing on the geometric mean definitions, average productivity growth in US agriculture in the period 1960–2004 lays in the 40.0%–50.0% range. Aggregate and individual TFP growth is mostly driven by technological progress, contributing around 30%. There are also mild efficiency gains, but scale, and input- and output-mix effects play a lesser role.

Fig. 2 depicts box-plots for the four productivity indices: Malmquist, Moorsteen-Bjurek, Fisher, and Törnqvist. Each box represents the interval between the first and third quartiles of the productivity change distribution, and the horizontal line represents the median. The dispersion of the distributions within these interquartile ranges appears to be rather small. Finally, the higher and lower whiskers signal values one and half times the interquartile range, while the dots represent outliers laying beyond those values. In the case of the Malmquist and Moorsteen-Bjurek indices, the highest and lowest values correspond to the state of Delaware (DE) and Oklahoma (OK), respectively. Note also that the price-based indices exhibit less variability and no outliers.

It is interesting to compare our results with those reported by the USDA (see Table 19, Indices of Total Factor Productivity by State, available at <https://www.ers.usda.gov/data-products/agricultural-productivity-in-the-us/>). The USDA Economic Research Service (ERS) publishes TFP measures based on a sophisticated system of farm production accounts. The measures are based on a translog transformation function which relates the growth rates of multiple outputs to the cost-share weighted growth rates of labor, capital, and intermediate inputs. See Ball et al. (1999) for a detailed description.

Table 2 displays correlations and distribution tests between TFP growth reported by the USDA and those obtained by the indices discussed in this paper. The results show a positive and significant correlation – Pearson's  $r$ , Spearman's  $\rho$ , and Kendall's  $\tau$  – between the USDA TFP index numbers and the Malmquist and Moorsteen-Bjurek index numbers, and a stronger correlation with the Fisher and Törnqvist index numbers. However, the  $t$  test on the equality of means and the Wilcoxon matched-pairs signed-rank test reject the null hypotheses that the means and the distributions of the USDA and this paper's index numbers are the same. The underlying reason is presumably that input cost shares obtained via the estimation procedure employed by USDA substantially differ from those obtained by DEA or conventional indices. It is, however, reassuring to see that the ranking of states tends to be preserved, as shown by the positive Spearman and Kendall rank correlation coefficients.

## 8. Conclusion

The **Total Factor Productivity Toolbox** allows the calculation and decomposition of the most popular productivity indices in an organized environment. The concept of distance function, implemented through non-parametric Data Envelopment Analysis, is the building block for the various decompositions. If only quantities are available, the toolbox calculates and decomposes TFP change according to the Malmquist and Moorsteen–Bjurek definitions. If also prices are available, classical TFP indices such as Fisher and Törnqvist can be calculated and decomposed.

Complex and lengthy decompositions involving panel data are rendered operational in a systematic way. The toolbox allows for many options concerning orientation, benchmark period, VRS or CRS, type of TFP index, and type of decomposition, so as to identify the contribution that different factors make to productivity change. We showed how to organize the data, use the available functions, and interpret the results. To illustrate the toolbox we calculated all the indices on a common example so that results can easily be compared. The toolbox is a valid self-contained MATLAB package for the measurement and decomposition of TFP change.

From the empirical findings, we draw some general conclusions:

- (i) Since the benchmark period matters empirically, we recommend researchers to use geometric means of base-period and comparison-period benchmark indices, unless there are compelling reasons that a particular benchmark should be used.
- (ii) The Malmquist and Moorsteen–Bjurek indices, based on quantities only, deliver higher productivity growth figures than the price-weighted and share-weighted indices which are based on quantities and (market) prices. To which extent this is due to the fact that the technologies were estimated by DEA in our empirical exercise remains to be seen.
- (iii) The results of price-weighted (for example, Fisher) and share-weighted (for example, Törnqvist) indices are numerically very close. Thus the choice between the two families of indices is immaterial.
- (iv) In our empirical example (local) technological change generally appeared to be the main component of TFP change.

Finally, since the code is freely available in an open source repository, users will benefit from the collaboration and review of the community. They can check and modify the code to adapt it to their own needs and extend it with new definitions.

## Acknowledgments

We are grateful to the editors Francisco Saldanha-da-Gama and Alice Smith, and to two anonymous reviewers for useful comments and suggestions. José L. Zofio acknowledges support from the [Spanish Ministry](#) for Economy and Competitiveness under grants [ECO2016-79650-P](#) and [MTM2016-79765-P](#). The authors acknowledge the financial support of the [University of Oviedo](#) through the project [PAPI-19-GR-2011-0026](#). The views expressed in this paper are those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.cor.2019.104853](https://doi.org/10.1016/j.cor.2019.104853)

## References

- Balk, B.M., 1998. *Industrial Price, Quantity, and Productivity Indices: The Micro-Economic Theory and an Application*. Springer US. [10.1007/978-1-4757-5454-4](https://doi.org/10.1007/978-1-4757-5454-4)
- Balk, B.M., 2001. Scale efficiency and productivity change. *J. Prod. Anal.* 15 (3), 159–183. doi:[10.1023/A:101117324278](https://doi.org/10.1023/A:101117324278).
- Balk, B.M., 2003. The residual: on monitoring and benchmarking firms, industries, and economies with respect to productivity. *J. Prod. Anal.* 20 (1), 5–47. doi:[10.1023/A:1024817024364](https://doi.org/10.1023/A:1024817024364).
- Balk, B. M., Zofio, J. L., 2018. The many decompositions of total factor productivity changeNo. ERS-2018-003-LIS. ERIM Report Series Research in Management. Erasmus Research Institute of Management. Erasmus University.
- Ball, V.E., Färe, R., Grosskopf, S., Nehring, R., 2001. Productivity of the u.s. agricultural sector: the case of undesirable outputs. In: Hulten, C.R., Dean, E.R., Harper, M.J. (Eds.), *New Developments in Productivity Analysis*. University of Chicago Press, Chicago, pp. 541–586.
- Ball, V.E., Gollop, F.M., Kelly-Hawke, A., Swinand, G.P., 1999. Patterns of state productivity growth in the u.s. farm sector: linking state and aggregate models. *Am. J. Agric. Econ.* 81 (1), 164–179.
- Banker, R., Charnes, A., Cooper, W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manag. Sci.* 30 (9), 1078–1092. doi:[10.1287/mnsc.30.9.1078](https://doi.org/10.1287/mnsc.30.9.1078).
- Bogotof, P., Otto, L., 2011. Benchmarking with DEA, SFA, and R. Springer-Verlag, New York doi:[10.1007/978-1-4419-7961-2](https://doi.org/10.1007/978-1-4419-7961-2).
- Caves, D.W., Christensen, L.R., Diewert, W.E., 1982. The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica* 50 (6), 1393–1414. doi:[10.2307/1913388](https://doi.org/10.2307/1913388).
- Charnes, A., Cooper, W., Rhodes, E., 1978. Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* 2 (6), 429–444. doi:[10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8).
- Coelli, T., 1998. A guide to deap v2.1: a data envelopment analysis (computer) programCEPA Working Paper 96/08. Centre for Efficiency and Productivity Analysis, University of Queensland.
- Coelli, T., 1999. Tfpip v1: Total factor productivity index (computer) programCentre for Efficiency and Productivity Analysis, University of Queensland.
- Cooper, W.W., Seiford, L.M., Tone, K., 2007. *Data Envelopment Analysis: a Comprehensive Text with Models, Applications, References and DEA-Solver Software*. NJ: Princeton University Press, Princeton doi:[10.1007/978-0-387-45283-8](https://doi.org/10.1007/978-0-387-45283-8).
- Dakpo, K.H., Desjeux, Y., Latruffe, L., 2018. Productivity: Indices of Productivity and Profitability Using Data Envelopment Analysis (DEA). R package version 1.1.0, <https://CRAN.R-project.org/package=productivity/>.
- Daraio, C., Kerstens, K., Nepomuceno, T., Sickles, R.C., 2019. Productivity and efficiency analysis software: an exploratory bibliographical survey of the options. *J. Econ. Surv.* 33, 85–100. doi:[10.1111/joes.12270](https://doi.org/10.1111/joes.12270).
- Econometric Software, I., 2014. LIMDEP11. Plainview, NY.
- Emrouznejad, A., Thanassoulis, E., 2011. Performance improvement management software: Pim-deasoft-v3.0 user guidePerformance Improvement Management Ltd, Coventry, U.K.
- Färe, R., Grosskopf, S., Lindgren, B., Roos, P., 1994. Productivity developments in Swedish hospitals: a malmquist output index approach. In: Charnes, A., Cooper, W.W., Lewin, A.Y., Seiford, L.M. (Eds.), *Data Envelopment Analysis: Theory, Methodology, and Applications*. Springer Netherlands, Dordrecht, pp. 253–272. doi:[10.1007/978-94-011-0637-5\\_13](https://doi.org/10.1007/978-94-011-0637-5_13).
- Färe, R., Grosskopf, S., Margaritis, D., 2008. Efficiency and productivity: Malmquist and more. In: Fried, H.O., Lovell, C.A.K., Schmidt, S.S. (Eds.), *The Measurement of Productive Efficiency and Productivity Change*. Oxford University Press, New York, pp. 522–621. doi:[10.1093/acprof:oso/9780195183528.003.0005](https://doi.org/10.1093/acprof:oso/9780195183528.003.0005).
- Färe, R., Grosskopf, S., Norris, M., Zhang, Z., 1994. Productivity growth, technical progress, and efficiency change in industrialized countries. *Am. Econ. Rev.* 84 (1), 66–83.
- Grifell-Tatjé, E., Lovell, C.A.K., 1995. A note on the malmquist productivity index. *Econ. Lett.* 47 (2), 169–175. doi:[10.1016/0165-1765\(94\)00497-P](https://doi.org/10.1016/0165-1765(94)00497-P).
- Lee, C., Lee, K.-R., Lim, B.-I., 2011. Malmquist productivity analysis using DEA frontier in STATAKorea National Defense University, Seoul, South Korea.
- Lovell, C.A.K., 2003. The decomposition of malmquist productivity indexes. *J. Prod. Anal.* 20 (3), 437–458. doi:[10.1023/A:1027312102834](https://doi.org/10.1023/A:1027312102834).
- Moorsteen, R.H., 1961. On measuring productive potential and relative efficiency. *Q. J. Econ.* 75 (3), 451–467. doi:[10.2307/1885133](https://doi.org/10.2307/1885133).
- O'Donnell, C. J., 2011. Dpin 3.0: A program for decomposing productivity index numbersCentre for Efficiency and Productivity Analysis, University of Queensland.
- O'Donnell, C.J., 2012. An aggregate quantity framework for measuring and decomposing productivity change. *J. Prod. Anal.* 38 (3), 255–272. doi:[10.1007/s11223-012-0275-1](https://doi.org/10.1007/s11223-012-0275-1).
- O'Donnell, C.J., 2018. *Productivity and Efficiency Analysis*. Springer Nature Singapore.
- OECD, 2001. *Measuring Productivity. Measurement of Aggregate and Industry Level Productivity Growth. Organization for Economic Co-operation and Development*, Paris doi:[10.1787/9789264194519-en](https://doi.org/10.1787/9789264194519-en).
- Ray, S.C., Desli, E., 1997. Productivity growth, technical progress, and efficiency change in industrialized countries: comment. *Am. Econ. Rev.* 87 (5), 1033–1039.
- Shehata, E. A. E., Mickael, S. K. A., 2015. Index: Stata module to estimate price, quantity, and value index numbersStatistical Software Components No. S458005. Department of Economics, Boston College, Boston, Mass., USA.
- Simar, L., Wilson, P.W., 2002. Non-parametric tests of returns to scale. *Eur. J. Oper. Res.* 139 (1), 115–132. doi:[10.1016/S0377-2217\(01\)00167-9](https://doi.org/10.1016/S0377-2217(01)00167-9).
- StataCorp, 2015. *Stata Statistical Software: Release 14*. StataCorp LP. College Station, TX.

- The MathWorks, Inc., 2017. MATLAB – *The Language of Technical Computing*, Version R2017b (9.3). Natick, Massachusetts.
- Zhu, J., 2014. *Quantitative Models for Performance Evaluation and Benchmarking*. Springer International Publishing AG, Cham, Switzerland doi:[10.1007/978-3-319-06647-9](https://doi.org/10.1007/978-3-319-06647-9).
- Zofio, J.L., 2007. Malmquist productivity index decompositions: a unifying framework. *Appl. Econ.* 39 (18), 2371–2387. doi:[10.1080/00036840600606260](https://doi.org/10.1080/00036840600606260).
- Zofio, J.L., Lovell, C.A.K., 2001. Graph efficiency and productivity measures: an application to us agriculture. *Appl. Econ.* 33 (11), 1433–1442. doi:[10.1080/00036840010009865](https://doi.org/10.1080/00036840010009865).