Measuring technical efficiency and total factor productivity change with undesirable outputs in Stata

Daoping Wang
School of Urban and Regional Science,
Shanghai University of Finance and Economics
Shanghai, China
daopingwang@live.sufe.edu.cn

Kerui Du
School of Management, Xiamen University
Xiamen, China
kerrydu@xmu.edu.cn

Ning Zhang
Institute of Blue and Green Development, Shandong University
Weihai, China
zn928@naver.com

Abstract. In this article, we introduce two user-written data envelopment analysis commands for measuring technical efficiency and productivity change in Stata. Over the last decades, an important theoretical progress of data envelopment analysis, a nonparametric method widely used for performance assessment of decision-making units, is the incorporating of undesirable outputs. Models with the ability to deal with undesirable outputs have been developed and applied in empirical studies for assessing the sustainability of decision-making units, and are getting more and more attention from researchers and managers. The `teddf` command developed in the present article allows users to measure technical efficiency, both radial and non-radial, when some outputs are undesirable. Technical efficiency measures are obtained by solving linear programming problems. The `gtfpch` command provides tools for measuring productivity change, e.g., the Malmquist-Luenberger index and the Luenberger indicator. We provide a brief overview of the nonparametric efficiency and productivity change measurement with the consideration of undesirable outputs, and we describe the syntax and options of the new commands. Examples are given to illustrate how to perform the technical efficiency and productivity analysis with the newly introduced commands.

Keywords: teddf, gtfpch, DEA, Malmquist-Luenberger index, Luenberger indicator, directional distance function, total fatcor productivity

1 Introduction

After the pioneering work of Farrell (1957), Debreu (1951), and Koopmans (1951), efficiency and productivity analysis have been widely used in empirical studies assessing the performance of decision-making units (DMUs) in terms of converting inputs into outputs. Among the parametric and nonparametric frontier models that have been developed in the field of efficiency analysis, Data Envelopment Analysis (DEA) has received plenty of attention for its no needing for prior information of the production function form and capability in multiple output technologies (Färe et al. 1985, 1994). Efficiency analysis based on DEA is usually assumed that inputs should be shrunk and outputs should be expanded. However, in the real world, outputs are not always desirable. In the case of undesirable outputs, they should be shrunk to improve efficiency. In the last decades, with the increasing demand for improving the sustainability of the economic society, scholars and managers gradually recognized that it is vital to consider undesirable output in efficiency and productivity analysis (Chung et al. 1997; Mahlberg and Sahoo 2011). Correspondingly, DEA models with the ability to deal with undesirable outputs have been developed and applied in empirical studies, e.g., (Zhou et al. 2012; Lin and Du 2015).

The estimation of nonparametric frontier models can be readily performed in Stata with some user-written commands. The `dea` command proposed in Ji and Lee (2010) provided a basic tool to estimate radial technical efficiency using the DEA technique in Stata. Badunenko and Mozharovskyi (2016) extended `dea` with five new commands that allow users to implement both radial and non-radial technical efficiency estimation, as well as statistical inference in nonparametric frontier models in Stata. Tauchmann (2012) introduced two commands, i.e., `orderm` and `orderalpha`, for implementing order-m, order-α, and free disposal hull efficiency analysis in Stata. These commands mentioned above, however, are limited in their capability for performing efficiency and productivity analysis with undesirable outputs.

Here, we introduce two user-written commands for measuring technical efficiency and productivity change with undesirable outputs in Stata. `teddf` estimates directional distance function with undesirable outputs for technical efficiency measurement. Both radial Debreu-Farrell and non-radial Russell measures can be calculated under different assumptions about the production technology, e.g., window, biennial, sequential, and global production technology. `gtfpch` measures total factor productivity (TFP) change with undesirable outputs using the Malmquist–Luenberger productivity index or the Luenberger indicator. The new commands open up the possibility to do efficiency and productivity analysis with undesirable outputs in Stata in an effortless way, and their results can directly feed to other Stata routines for further analysis.

The remainder of this article unfolds as follows: section 2 provides a brief overview of the nonparametric efficiency and productivity change measurement with the consideration of undesirable outputs; sections 3 and 4 contain the syntax and explain the options of `teddf` and `gtfpch`, respectively; section 5 presents a hypothetical example to show the usage of the two commands; and section 6 concludes the article.

2 The model

In this section, we provide a brief overview of the nonparametric efficiency and productivity measurement with the consideration of undesirable outputs. The radial and non-radial directional distance function (DDF) will be introduced firstly, followed by the description of the measurement of technical efficiency and TFP change using DDFs. The exposition here is only introductory. For more details, please refer to the cited works.

2.1 Directional distance function

Consider a production unit transforming a vector of nonnegative inputs into a vector of nonnegative desirable outputs and a vector of by-products (undesirable outputs) such as pollution, subject to the constraint imposed by a fixed technology. For the production technology, inputs and desirable outputs are supposed to be strongly disposable. The undesirable outputs are assumed to be weakly disposable, which indicates that the decrease of undesirable outputs is not free but will lead to a deduction in desirable outputs. If we denote the input, desirable output, and undesirable outputs as $\boldsymbol{x} \in \Re_+^N$ $\boldsymbol{y} \in \Re_+^M$, and $\boldsymbol{b} \in \Re_+^H$, respectively. This production technology can be characterized by the technology set as

$$T = \{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}) : \boldsymbol{x} \text{ can produce } (\boldsymbol{y}, \boldsymbol{b})\}. \tag{1}$$

Then, following Chung et al. (1997), the radial DDF is defined as

$$D_r(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}; \boldsymbol{g}) = \sup\{\beta : ((\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}) + \beta \boldsymbol{g}) \in T\} \tag{2}$$

where $\boldsymbol{g} = (\boldsymbol{g}_x, \boldsymbol{g}_y, \boldsymbol{g}_b) \in \Re_-^N \times \Re_+^M \times \Re_-^H$ is a preassigned nonzero vector, specifying the direction in which the distance between the data point, $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b})$, and the production frontier is measured.

Equation (2) gives out the most general form of the DDF. One can define the distance between the DMU and the production frontier in a specific direction by setting different $\boldsymbol{g}$. By way of illustration, we consider the cases of $\boldsymbol{g}_1 = (\boldsymbol{0}, \boldsymbol{y}, \boldsymbol{0})$, $\boldsymbol{g}_2 = (\boldsymbol{0}, \boldsymbol{0}, -\boldsymbol{b})$, $\boldsymbol{g}_3 = (\boldsymbol{0}, \boldsymbol{y}, -\boldsymbol{b})$, which is widely used in literature. Figure 1 presents hypothetical one-desirable output (e.g., GDP) one-undesirable output (e.g., CO_2) production processes. Conceptually, in Fig. 1, $\overline{AB}$ and $\overline{AC}$ represent the distance when the direction is $\boldsymbol{g}_1 = (\boldsymbol{0}, \boldsymbol{y}, \boldsymbol{0})$ and $\boldsymbol{g}_2 = (\boldsymbol{0}, \boldsymbol{0}, -\boldsymbol{b})$, respectively. The former focuses on economic prosperity, while the latter focuses on environmental protection. Similarly, $\overline{AD}$ is the distance when the direction is $\boldsymbol{g}_3 = (\boldsymbol{0}, \boldsymbol{y}, -\boldsymbol{b})$, which describes the maximum increase of desirable output while simultaneously reducing the undesirable output along the direction $(\boldsymbol{y}, -\boldsymbol{b})$. Intuitively, the smaller the distance, the closer the DMU is next to the production frontier, and the distance is 0 for the DMU which operates on the production frontier.

The radial measure expands (shrinks) all outputs or/and inputs proportionally until the production frontier is reached. At the reached frontier point, some

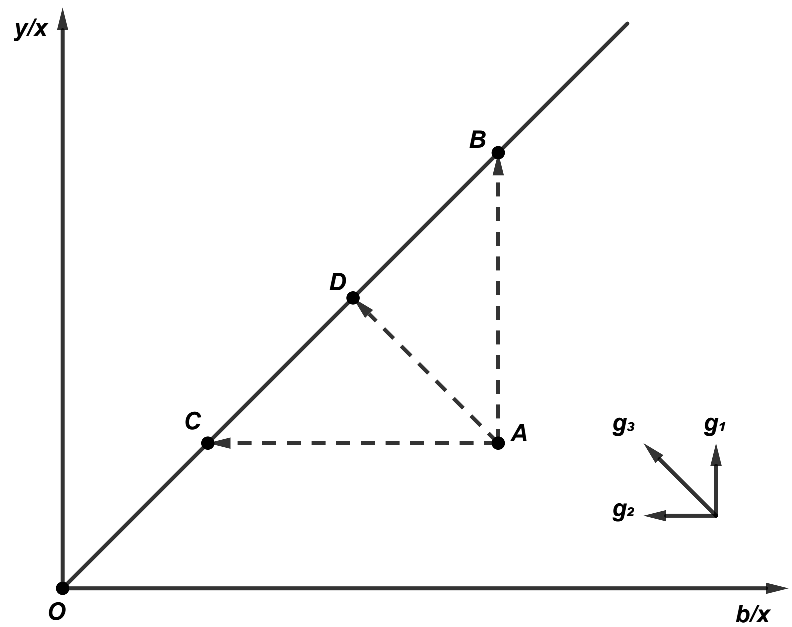

Figure 1: A graphical illustration of directional distance functions

but not all outputs (inputs) can be expanded (shrunk) while remaining feasible. If such a possibility is available for a given decision-making unit for some outputs (inputs), then the reference point is said to have slacks in outputs (inputs). Non-radial measures, i.e., the Russell measure, accommodate such slacks (Chambers 2002; Färe and Grosskopf 2010; Zhou et al. 2012).

The non-radial DDF is defined as

$$D_{nr}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}; \boldsymbol{g}) = \sup\{\boldsymbol{w}^T \boldsymbol{\beta} : ((\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}) + diag(\boldsymbol{\beta}) \cdot \boldsymbol{g}) \in T\} \tag{3}$$

where $\boldsymbol{w}$ denotes a normalized weight vector that is relevant to the numbers of inputs and outputs, and $\boldsymbol{\beta} = (\boldsymbol{\beta}_x, \boldsymbol{\beta}_y, \boldsymbol{\beta}_b) \in \Re^N \times \Re^M \times \Re^H$ denotes the vector of the scaling factors. Clearly, the non-radial DDF measure allows the inputs and outputs to be adjusted non-proportionally. Compared with the radial measure in Fig.1, instead of using a fixed point, e.g., B, C, or D, as the reference point, if the non-radial directional distance function is used, the reference point would be located at any point on the production frontier.

2.2 Measurement of technical efficiency

To estimate the DDF measure of technical efficiency using the nonparametric technique, the production technology set is derived from observed data. Typically, for the cross-sectional data, the production technology set with the assumption of constant return to scale (CRS) is constructed as

$$T = \left\{ (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}) : \sum_{j=1}^{J} \lambda_j \boldsymbol{x}_j \leq \boldsymbol{x}, \sum_{j=1}^{J} \lambda_j \boldsymbol{y}_j \geq \boldsymbol{y}, \sum_{j=1}^{J} \lambda_j \boldsymbol{b}_j = \boldsymbol{b}, \boldsymbol{\lambda} \geq 0 \right\}. \tag{4}$$

For variable returns to scale (VRS), $\sum_{j=1}^{J} \lambda_j = 1$ is added to the above equation. That is,

$$T = \left\{ (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}) : \sum_{j=1}^{J} \lambda_j \boldsymbol{x}_j \leq \boldsymbol{x}, \sum_{j=1}^{J} \lambda_j \boldsymbol{y}_j \geq \boldsymbol{y}, \sum_{j=1}^{J} \lambda_j \boldsymbol{b}_j = \boldsymbol{b}, \boldsymbol{\lambda} \geq 0, \sum_{j=1}^{J} \lambda_j = 1 \right\}. \tag{5}$$

In the panel data context, the time-series dimension can provide more information on the production technology. Researchers have proposed different types of production technology sets such as global, sequential, window, biennial, and contemporaneous production technology. The production technology set at the time t as follows:

$$T(t) = \left\{ (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}) : \sum_{\tau \in \Gamma_t} \sum_{j=1}^{J} \lambda_{j\tau} \boldsymbol{x}_{j\tau} \leq \boldsymbol{x}, \sum_{\tau \in \Gamma_t} \sum_{j=1}^{J} \lambda_{j\tau} \boldsymbol{y}_{j\tau} \geq \boldsymbol{y}, \sum_{\tau \in \Gamma_t} \sum_{j=1}^{J} \lambda_{j\tau} \boldsymbol{b}_{j\tau} = \boldsymbol{b}, \boldsymbol{\lambda} \geq 0 \right\}. \tag{6}$$

The time range, $\tau \in \Gamma_t$, for different types of production technology set are shown in Fig.2. In the global production technology, $\tau \in \Gamma_t$ is expressed as $\tau \leq t_{max}$, where t_{max} is the last period in the sample. In the sequential production technology, $\tau \in \Gamma_t$ is expressed as $\tau \leq t$. In the window production technology, $\tau \in \Gamma_t$ is expressed as $t - h \leq \tau \leq t + h$, where h is the bandwidth. In the biennial production technology, $\tau \in \Gamma_t$ is expressed as $t \leq \tau \leq t + 1$. In the contemporaneous production technology, $\tau \in \Gamma_t$ is expressed as $\tau = t$.

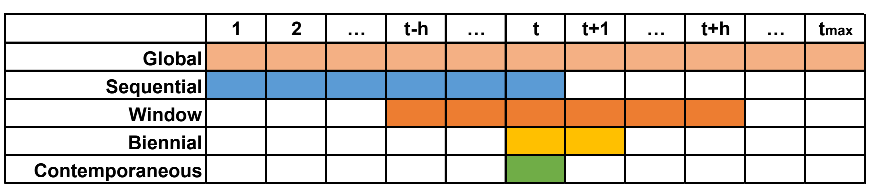

Figure 2: Timing assumption of different type of production technology sets

Then, the radial DDF measure of inefficiency under the CRS assumption can be estimated by solving the following linear programming problem,

$$
\begin{aligned}
D_r(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}; \boldsymbol{g}) = & \max_{\beta, \boldsymbol{\lambda}} \beta \\
& \text{s.t.} \sum_{j=1}^{J} \lambda_j \boldsymbol{x}_j \leq \boldsymbol{x} + \beta \boldsymbol{g}_x, \\
& \sum_{j=1}^{J} \lambda_j \boldsymbol{y}_j \geq \boldsymbol{y} + \beta \boldsymbol{g}_y, \\
& \sum_{j=1}^{J} \lambda_j \boldsymbol{b}_j = \boldsymbol{b} + \beta \boldsymbol{g}_b, \\
& \lambda_j \geq 0, j = 1, ..., J.
\end{aligned} \tag{7}
$$

This technology set is based on the assumption of constant return to scale (CRS). For VRS assumption, $\sum_{j=1}^{J} \lambda_j = 1$ is added to above constraints.

In Eq.(7), the left-hand side of the constraints construct the production frontier using the convex hull of the observation data and the right-hand side allows the assessed DMU to adjust the inputs (x), the desirable outputs (y) and undesirable outputs (b) alongside the direction of (g_x, g_y, g_b). The directional distance function seeks to maximize the reduction of inputs and undesirable outputs and the expansion of the desirable output in means of $(x + \beta g_x, y + \beta g_y, b + \beta g_b)$ given the production technology.

Similarly, the non-radial DDF measure of inefficiency considering the undesirable outputs can be obtained by solving the following linear programming problem.

$$
\begin{aligned}
D_{nr}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}; \boldsymbol{g}) = & \max_{\boldsymbol{\beta}, \boldsymbol{\lambda}} \boldsymbol{w}^T \boldsymbol{\beta} \\
& \text{s.t.} \sum_{j=1}^{J} \lambda_j \boldsymbol{x}_j \leq \boldsymbol{x} + diag(\boldsymbol{\beta}_x) \cdot \boldsymbol{g}_x, \\
& \sum_{j=1}^{J} \lambda_j \boldsymbol{y}_j \geq \boldsymbol{y} + diag(\boldsymbol{\beta}_y) \cdot \boldsymbol{g}_y, \\
& \sum_{j=1}^{J} \lambda_j \boldsymbol{b}_j = \boldsymbol{b} + diag(\boldsymbol{\beta}_b) \cdot \boldsymbol{g}_b, \\
& \boldsymbol{\beta} \geq 0; \lambda_j \geq 0, j = 1, ..., J.
\end{aligned} \tag{8}
$$

For variable returns to scale (VRS) assumption, $\sum_{j=1}^{J} \lambda_j = 1$ is added to above constraints.

Unlike the directional distance function shown in Eq.(7), the non-radial directional distance function allows each component of inputs, desirable outputs, and undesirable outputs to adjust in varying proportions.The non-radial DDF is the maximum weighted sum of the adjustment components (β) such that

$(\beta_x \cdot g_x, diag(\beta_y) \cdot g_y, diag(\beta_b) \cdot g_b)$ can be produced given the production technology.

For the panel data case, the radial DDF measure of inefficiency under the CRS assumption can be estimated by solving the following linear programming problem,

$$\begin{aligned} D_r(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}; \boldsymbol{g}) = &\max_{\beta, \boldsymbol{\lambda}} \beta \\ &\text{s.t.} \sum_{\tau \in \Gamma_t} \sum_{j=1}^{J} \lambda_j \boldsymbol{x}_{j\tau} \leq \boldsymbol{x} + \beta \boldsymbol{g}_x, \\ &\sum_{\tau \in \Gamma_t} \sum_{j=1}^{J} \lambda_{j\tau} \boldsymbol{y}_{j\tau} \geq \boldsymbol{y} + \beta \boldsymbol{g}_y, \\ &\sum_{\tau \in \Gamma_t} \sum_{j=1}^{J} \lambda_{j\tau} \boldsymbol{b}_{j\tau} = \boldsymbol{b} + \beta \boldsymbol{g}_b, \\ &\lambda_{j\tau} \geq 0, j = 1, ..., J. \end{aligned} \tag{9}$$

Similarly, for the panel data case, the non-redial DDF measure of inefficiency can be estimated by solving the following linear programming problem,

$$\begin{aligned} D_{nr}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}; \boldsymbol{g}) = &\max_{\boldsymbol{\beta}, \boldsymbol{\lambda}} \boldsymbol{w}^T \boldsymbol{\beta} \\ &\text{s.t.} \sum_{\tau \in \Gamma_t} \sum_{j=1}^{J} \lambda_j \boldsymbol{x}_{j\tau} \leq \boldsymbol{x} + diag(\boldsymbol{\beta}_x) \cdot \boldsymbol{g}_x, \\ &\sum_{\tau \in \Gamma_t} \sum_{j=1}^{J} \lambda_j \tau \boldsymbol{y}_{j\tau} \geq \boldsymbol{y} + diag(\boldsymbol{\beta}_y) \cdot \boldsymbol{g}_y, \\ &\sum_{\tau \in \Gamma_t} \sum_{j=1}^{J} \lambda_j \tau \boldsymbol{b}_{j\tau} = \boldsymbol{b} + diag(\boldsymbol{\beta}_b) \cdot \boldsymbol{g}_b, \\ &\boldsymbol{\beta} \geq 0; \lambda_{j\tau} \geq 0, j = 1, ..., J. \end{aligned} \tag{10}$$

2.3 Measurement of total factor productivity change

The measurement of productivity change has traditionally focused on measuring marketable (desirable) outputs of DMUs relative to paid factors of production. This approach, which typically ignores the production of by-products such as pollution, can yield biased measures of productivity growth (Chung et al. 1997). For example, firms in industries that face environmental regulations would typically find that their productivity is adversely affected since the costs of abatement capital would typically be included on the input side, but no account would be made of the reduction in pollutants on the output side.

Chung et al. (1997) has introduced a productivity index based on the radial DDF measure, called the Malmquist-Luenberger productivity index, which credits the reduction of undesirable outputs, e.g., pollution, while simultaneously crediting increases in desirable outputs. Considering two adjacent periods, denoted as s and t. respectively. If we choose the direction to be $\boldsymbol{g} = (\boldsymbol{0}, \boldsymbol{y}, -\boldsymbol{b})$, the output-oriented Malmquist-Luenberger productivity index with undesirable outputs is defined as

$$ML = \left[\frac{1 + D_r^s(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g})}{1 + D_r^s(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g})} \times \frac{1 + D_r^t(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g})}{1 + D_r^t(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g})}\right]^{1/2}. \tag{11}$$

To avoid an arbitrary choice between base years, an geometric mean of a fraction-based Malmquist-Luenberger productivity index in base year t (first fraction) and s (second fraction) has been taken. The Malmquist–Luenberger measure indicates productivity improvements if their values are greater than one and decreases in productivity if the values are less than one.

The Malmquist-Luenberger productivity index can be decomposed into two components (Chung et al. 1997), one accounting for efficiency change (*MLEFFCH*), and one measuring technology change (*MLTECH*):

$$MLEFFCH = \frac{1 + D_r^s(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g})}{1 + D_r^t(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g})} \tag{12}$$

and,

$$MLTECH = \left[\frac{1 + D_r^t(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g})}{1 + D_r^s(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g})} \times \frac{1 + D_r^t(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g})}{1 + D_r^s(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g})}\right]^{1/2}. \tag{13}$$

Based on the pioneering work in Chambers (2002), another productivity measure called the Luenberger productivity indicator is also widely used to account for productivity change. The Luenberger productivity indicator based on radial DDF measures is defined as

$$\begin{aligned} L = &\left[(D_r^t(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g}) - D_r^t(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g})\right] \times \frac{1}{2} \\ &+ \left[D_r^s(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g}) - D_r^s(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g}))\right] \times \frac{1}{2}. \end{aligned} \tag{14}$$

Again, to avoid an arbitrary choice between base years, an arithmetic mean of a difference based Luenberger productivity index in base year t (first difference) and s (second difference) has been taken. Productivity improvements are indicated by positive values and declines by negative values.

In the spirit of decomposition of Malmquist–Luenberger productivity index, the Luenberger productivity indicator based on radial DDFs can also be decomposed into two component measures, i.e., an efficiency change component (Mahlberg and Sahoo 2011),

$$LEFFCH = D_r^s(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g}) - D_r^t(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g}) \tag{15}$$

and a technical change component,

$$\begin{aligned} LTECH = &\left[D_r^t(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g}) - D_r^s(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g})\right] \times \frac{1}{2} \\ &+ \left[(D_r^t(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g}) - D_r^s(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g})\right] \times \frac{1}{2}. \end{aligned} \tag{16}$$

The Luenberger indicator based on radial DDFs is expressed as the sum of *LEFFCH* and *LTECH*. *LEFFCH* captures the average gain/loss due to the difference in technical efficiency from period s to period t. *LTECH* captures the average gain/loss due to the shift in technology from period s to period t.

Like any radial measure of efficiency estimated using DEA technologies, DDF overestimates the efficiency of a firm when there are non-zero slacks that remain in the constraints after the full radial efficiency is achieved. To account for these slacks, Färe and Grosskopf (2010) proposed a slacks-based measure of efficiency based on non-radial directional distance function. Another type of Luenberger indicator, called the non-proportional Luenberger indicator, can be constructed based on the non-radial DDFs (Mahlberg and Sahoo 2011).

The Luenberger productivity indicator based on non-radial DDFs is defined as

$$\begin{aligned} L = & \left[(D^t_{nr}(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g}) - D^t_{nr}(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g})\right] \times \frac{1}{2} \\ & + \left[D^s_{nr}(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g}) - D^s_{nr}(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g}))\right] \times \frac{1}{2}. \end{aligned} \tag{17}$$

The non-proportional Luenberger productivity indicator can also be decomposed into two parts (Mahlberg and Sahoo 2011). An efficiency change component,

$$LEFFCH = D^s_{nr}(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g}) - D^t_{nr}(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g}) \tag{18}$$

and a technical change component,

$$\begin{aligned} LTECH = & \left[D^t_{nr}(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g}) - D^s_{nr}(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{b}^t; \boldsymbol{g})\right] \times \frac{1}{2} \\ & + \left[(D^t_{nr}(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g}) - D^s_{nr}(\boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{b}^s; \boldsymbol{g})\right] \times \frac{1}{2}. \end{aligned} \tag{19}$$

2.4 Statistical inference for technical efficiency and productivity index

Efficiency and productivity analysis are widely applied in benchmarking(relative performance evaluations). The models introduced above are based on the DEA methods which are typically considered to be deterministic. Specifically, the efficiency/productivity is measured relative to the estimated production frontiers constructed by sample observations. Consequently, the measures of efficiency/productivity might be sensitive to the sampling variations. In view of this, researchers have devoted themselves to exploring the statistical properties of the DEA-type estimators. For instance, Banker (1993) and Kneip et al. (1998) established consistency and convergence rates of DEA efficiency estimators. Kneip et al. (2008) derived the asymptotic distribution of Farrell measure of technical efficiency (one of the radial DEA estimators) in cases with multiple inputs and outputs. Generally speaking, many DEA-type efficiency estimators have been proposed. But few have been known on their asymptotic distribution.

To implement statistical inference for DEA-type efficiency estimators, Simar and Wilson (1998) proposed a smoothed bootstrapping procedure. But the consistency of the smooth bootstrapping method has not been proved. Based on the asymptotic theorems developed in Kneip et al. (2008), they presented another two consistent bootstrapping procedures for Farrell measure of technical efficiency: the subsampling approach and the double-smooth bootstrap. Simar et al. (2012) showed that the directional distance function estimators shared the known properties of the traditional radial DEA estimators and adapted the subsampling approach and the double-smooth bootstrap to this context. In the context of non-radial DEA estimators, Badunenko and Mozharovskyi (2020) proposed a bootstrap method for Russell measures of technical efficiency. Badunenko and Mozharovskyi (2016) incorporated the bootstrap procedures in their Stata commands `teradialbc` and `tenonradialbc`. It is worth pointing out that the bootstrap methods mentioned above mainly focus on the cross-sectional data. In the panel data cases, there are some difficulties in applying the smooth/double-smooth bootstrapping methods. Because considering the possibility of temporal correlation, they require nonparametric estimation of a high dimensional density which might suffer from the curse of dimensionality. On the contrary, the subsampling approach can be easily adapted to accommodate the panel data structure by subsampling with clusters. Thus, we incorporate the subsampling approach in our Stata command (`teddf`) for statistical inference.

Regarding the statistical inference for DEA-based productivity indexes, Simar and W. Wilson (2019) established the asymptotic theorems for nonparametric Malmquist indices. Simar and Wilson (1999) proposed a bootstrap estimation procedure for obtaining confidence intervals for Malmquist indices of productivity and their decompositions. But until now, the statistical properties of the Malmquist-Luenberger productivity index and the Luenberger productivity indicator are still unknown. Intuitively, the subsampling approach can be adapted to these contexts with the knowledge of the convergence rates. Nevertheless, it is still an open issue.

3 The teddf command

`teddf` estimates directional distance function with undesirable outputs for technical efficiency measurement.

3.1 Syntax

`teddf` *Xvarlist* `=` *Yvarlist*`:`*Bvarlist* [*if*] [*in*], `dmu(`*varname*`)` [
`time(`*varname*`)` `gx(`*varlist*`)` `gy(`*varlist*`)` `gb(`*varlist*`)` `nonradial`
`wmat(name)` `vrs` `rf(`*varname*`)` `window(#)` `biennial` `sequential`
`global` `brep(#)` `alpha(real)` `tol(real)` `maxiter(#)`
`saving(filename[,replace])` `frame(framename)` `nodots` `noprint`
`nocheck`]

3.2 Options

`dmu(`*varname*`)` specifies names of DMUs. It is required.

`time(`*varname*`)` specifies the time variable for panel data.

`gx(`*varlist*`)` specifies direction components for input adjustment. The order of variables specified in gx() should as the same in *Xvarlist*. By default, gx takes the opposite of *Xvarlist*.

`gy(`*varlist*`)` specifies direction components for desirable output adjustment. The order of variables specified in gy() should as the same in *Yvarlist*. By default, gy takes *Yvarlist*.

`gb(`*varlist*`)` specifies direction components for undesirable output adjustment. The order of variables specified in gb() should as the same in *Bvarlist*. By default, gb takes the opposite of *Bvarlist*.

`nonradial` specifies using the nonradial directional distance measure.

`wmat(name)` specifies a weight matrix for adjustment of input and output variable for the nonradial directional distance measure. The default is wmatrix= (1,...,1).

`vrs` specifies production technology with variable returns to scale. By default, production technology with constant returns to scale is assumed.

`rf(`*varname*`)` specifies the indicator variable that defines which data points of outputs and inputs form the technology reference set.

`window(#)` specifies using window production technology with the #-period bandwidth.

`biennial` specifies using biennial production technology.

`sequential` specifies using sequential production technology.

`global` specifies using global production technology.

`brep(#)` specifies the number of bootstrap replications. The default is brep(0) specifying performing the estimator without bootstrap. Typically, it requires 1,000 or more replications for bootstrap DEA methods.

`alpha(real)` sets the size of the subsample bootstrap. By default, alpha(0.7) indicates subsampling $N^{0.7}$ observations out of the N original reference observations.

`tol(real)` specifies the convergence-criterion tolerance for LinearProgram(). The default value of tol is 1e-8.

`maxiter(#)` specifies the maximum number of iterations for LinearProgram(). The default value of maxiter is 16000.

`saving(filename[,replace])` specifies a filename to store the results.

`frame(name)` specifies a framename to stroe the results.

`nodots` suppress iteration dots.

`noprint` suppress suppress display of the results.

`nocheck` suppress checking for new version. It is suggested to be used for saving time when internet connection is unavailable.

4 The gtfpch command

`gtfpch` measures total factor productivity change with undesirable outputs using Malmquist–Luenberger productivity index or Luenberger indicator.

4.1 Syntax

`gtfpch` *Xvarlist* `=` *Yvarlist*`:`*Bvarlist* [*if*] [*in*] `,` [`dmu(`*varname*`)`
`luenberger ort(`*string*`) gx(`*varlist*`) gy(`*varlist*`) gb(`*varlist*`) nonradial`
`wmat(name) window(#) biennial sequential global fgnz rd`
`tol(real) maxiter(#) saving(filename[,replace])`
`frame(framename) noprint nocheck`]

4.2 Options

`dmu(`*varname*`)` specifies names of DMUs.

`luenberger` specifies estimating Luenberger productivity indicator. The default is Malmquist–Luenberger productivity index based on the radial directional distance function.

`ort(`*string*`)` specifies the oriention. The default is ort(input), meaning the input oriented productivity index/indicator. ort(output) means the output oriented productivity index/indicator. ort(hybrid) means the hybrid-direction productivity index/indicator.

`gx(`*varlist*`)` specifies direction components for input adjustment. The order of variables specified in gx() should as the same in *Xvarlist*. By default, gx=(0,..,0) for ort(output); gx=-*Xvarlist* for ort(input) and ort(hybrid).

`gy(`*varlist*`)` specifies direction components for desirable output adjustment. The order of variables specified in gy() should as the same in *Yvarlist*. By default, gy=(0,..,0) for ort(input); gy=*Yvarlist* for ort(output) and ort(hybrid).

`gb(`*varlist*`)` specifies direction components for undesirable output adjustment. The order of variables specified in gb() should as the same in *Bvarlist*. By default, gb=(0,..,0) for ort(input); gb=-*Bvarlist* for ort(output) and ort(hybrid).

`nonradial` specifies using the non-radial directional distance measure.

`wmat(name)` specifies a weight matrix for adjustment of input and output variable for the nonradial directional distance measure.

`window(#)` specifies using window production technology with the #-period bandwidth.

`biennial` specifies using biennial production technology.

`sequential` specifies using sequential production technology.

`global` specifies using global production technology.

`fgnz` specifies specifies decomposing TFP change following the spirit of Färe et al. (1994) method.

`rd` specifies decomposing TFP change following the spirit of Ray and Desli (1997) method.

`tol(real)` specifies the convergence-criterion tolerance for LinearProgram(). The default value of tol is 1e-8.

`maxiter(#)` specifies the maximum number of iterations for LinearProgram(). The default value of maxiter is 16000.

`saving(filename[,replace])` specifies a file name to store the results.

`frame(name)` specifies a frame name to store the results.
`noprint` suppress suppress display of the results.
`nocheck` suppress checking for new version. It is suggested to be used for saving time when internet connection is unavailable.

5 Example

To exemplify the use of the commands described above, we use an input-output data set of China's provinces for the period of 2013-2015 which is obtained from a recent publication, Yan et al. (2020). The dataset includes three input variables (capital, labor, and energy), one desirable output (real GDP), and one undesirable output (CO_2 emissions). The data are described as follows.

```
.
. use example.dta

. describe

Contains data from example.dta
  obs:            90
 vars:             7                          6 Aug 2020 12:12
--------------------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
--------------------------------------------------------------------------------------------
Province        str12   %12s                  province name
year            int     %10.0g                year
K               float   %9.0g                 capital stock (in 100 million 1997 CNY)
L               double  %10.0g                employment (in 10 thousand persons)
E               double  %10.0g                energy consumption (in million tons of standard coal)
Y               float   %9.0g                 real GDP (in 100 million 1997 CNY)
CO2             float   %15.1f                carbon dioxide emission (in kg)
--------------------------------------------------------------------------------------------
Sorted by:
.
```

5.1 Application of teddf

The estimation of the directional distance function model proposed by Chung et al. (1997) as follows. The corresponding results are displayed below the executed command. The Dv variable stores the values of the directional distance function of the DMUs. Note that the sav(ex.teddf.result) option saves the results in a new data file named ex.teddf.result.dta.

```
.
. teddf K L= Y: CO2, dmu(Province) time(year) sav(ex.teddf.result,replace)
 The directional vector is (-K -L Y -CO2)

  Directional Distance Function Results:
    (Row: Row # in the original data; Dv: Estimated value of  DDF.)

     +----------------------------------+
     | Row      Province   year      Dv |
     |----------------------------------|
  1. |   1         Anhui   2013  0.2917 |
  2. |   2         Anhui   2014  0.3589 |
```

```
  3. |   3          Anhui   2015    0.3735 |
  4. |   4        Beijing   2013   -0.0000 |
  5. |   5        Beijing   2014   -0.0000 |
  6. |   6        Beijing   2015   -0.0000 |
                      ...
                      ...
                      ...
 85. |  82         Yunnan   2013    0.2680 |
 86. |  83         Yunnan   2014    0.3446 |
 87. |  84         Yunnan   2015    0.3416 |
 88. |  85       Zhejiang   2013    0.1197 |
 89. |  86       Zhejiang   2014    0.1540 |
 90. |  87       Zhejiang   2015    0.1732 |
     +-------------------------------------+

Note: Missing value indicates infeasible problem.
(note: file ex.teddf.result.dta not found)
file ex.teddf.result.dta saved

Estimated Results are saved in ex.teddf.result.dta.

.
```

To customize the directional vector,

```
.
. gen gK=0

. gen gL=0

. gen gY=Y

. gen gCO2=-CO2

. teddf K L= Y: CO2, dmu(Province) time(year) gx(gK gL) gy(gY) gb(gCO2) sav(ex.teddf.direction.result,repla
 The directional vector is (gK gL gY gCO2)

  Directional Distance Function Results:
    (Row: Row # in the original data; Dv: Estimated value of  DDF.)

     +-------------------------------------+
     | Row       Province   year        Dv |
     |-------------------------------------|
  1. |   1          Anhui   2013    0.4024 |
  2. |   2          Anhui   2014    0.4515 |
  3. |   3          Anhui   2015    0.5049 |
  4. |   4        Beijing   2013   -0.0000 |
  5. |   5        Beijing   2014   -0.0000 |
  6. |   6        Beijing   2015   -0.0000 |
                      ...
                      ...
                      ...
 85. |  82         Yunnan   2013    0.4267 |
 86. |  83         Yunnan   2014    0.4428 |
 87. |  84         Yunnan   2015    0.4648 |
 88. |  85       Zhejiang   2013    0.1507 |
 89. |  86       Zhejiang   2014    0.1891 |
 90. |  87       Zhejiang   2015    0.2265 |
     +-------------------------------------+

Note: Missing value indicates infeasible problem.
(note: file ex.teddf.direction.result.dta not found)
file ex.teddf.direction.result.dta saved

Estimated Results are saved in ex.teddf.direction.result.dta.

.
```

Additionally, we show an application of **teddf** to estimate the non-radial directional distance function model as follows. The Dv variable stores the values of the non-radial directional distance function of the DMUs. B_K, B_L, B_CO2, and B_Y variables store the reduction proportion of inputs (K,L) and undesirable outputs ($CO2$), and the expansion proportion of desirable output (Y), respectively.

```
.
. teddf K L= Y: CO2, dmu(Province) time(year) nonr sav(ex.teddf.nonr.result,replace)
 The weight vector is (1 1 1 1)
 The directional vector is (-K -L Y -CO2)

 Non-raidal Directional Distance Function Results:
    (Row: Row # in the original data; Dv: Estimated value of Non-raidal DDF.)
```

	Row	Province	year	Dv	B_K	B_L	B_Y	B_CO2
1.	1	Anhui	2013	1.6710	0.4594	0.7225	0.0000	0.4890
2.	2	Anhui	2014	1.7823	0.5293	0.7198	0.0000	0.5331
3.	3	Anhui	2015	1.8210	0.5827	0.7181	0.0000	0.5202
4.	4	Beijing	2013	0.0000	0.0000	0.0000	0.0000	0.0000
5.	5	Beijing	2014	0.0000	0.0000	0.0000	0.0000	0.0000
6.	6	Beijing	2015	0.0000	0.0000	0.0000	0.0000	0.0000
				...				
				...				
				...				
85.	82	Yunnan	2013	1.7617	0.4480	0.7814	0.0000	0.5323
86.	83	Yunnan	2014	1.8300	0.5165	0.7834	0.0000	0.5301
87.	84	Yunnan	2015	1.8184	0.5696	0.7790	0.0000	0.4698
88.	85	Zhejiang	2013	0.8887	0.2696	0.4386	0.0000	0.1805
89.	86	Zhejiang	2014	1.0078	0.3364	0.4375	0.0000	0.2340
90.	87	Zhejiang	2015	1.0589	0.3912	0.4368	0.0000	0.2309

```
Note: Missing value indicates infeasible problem.
(note: file ex.teddf.nonr.result.dta not found)
file ex.teddf.nonr.result.dta saved
Estimated Results are saved in ex.teddf.nonr.result.dta.
.
```

To customize the weight matrix,

```
.
. mat wmatrix=(0.5,0.5,1,1)
. teddf K L= Y: CO2, dmu(Province) time(year) nonr wmat(wmatrix) sav(ex.teddf.nonr.weight.result,replace)
 The weight vector is (.5 .5 1 1)
 The directional vector is (-K -L Y -CO2)

 Non-raidal Directional Distance Function Results:
    (Row: Row # in the original data; Dv: Estimated value of Non-raidal DDF.)
```

Row	Province	year	Dv	B_K	B_L	B_Y	B_CO2

```
  1. | 1           Anhui   2013   1.1480   0.0000   0.4867   0.8499   0.0548 |
  2. | 2           Anhui   2014   1.3351   0.0000   0.4047   1.1247   0.0081 |
  3. | 3           Anhui   2015   1.3577   0.0000   0.3305   1.1924   0.0000 |
  4. | 4         Beijing   2013   0.0000   0.0000   0.0000   0.0000   0.0000 |
  5. | 5         Beijing   2014   0.0000   0.0000   0.0000   0.0000   0.0000 |
  6. | 6         Beijing   2015   0.0000   0.0000   0.0000   0.0000   0.0000 |
                                  ...
                                  ...
                                  ...
 85. | 82         Yunnan   2013   1.2663   0.0000   0.6040   0.8117   0.1526 |
 86. | 83         Yunnan   2014   1.3723   0.0000   0.5520   1.0682   0.0281 |
 87. | 84         Yunnan   2015   1.2843   0.0000   0.4933   1.0377   0.0000 |
 88. | 85       Zhejiang   2013   0.5346   0.2696   0.4386   0.0000   0.1805 |
 89. | 86       Zhejiang   2014   0.6209   0.3364   0.4375   0.0000   0.2340 |
 90. | 87       Zhejiang   2015   0.6449   0.3912   0.4368   0.0000   0.2309 |

Note: Missing value indicates infeasible problem.
(note: file ex.teddf.nonr.weight.result.dta not found)
file ex.teddf.nonr.weight.result.dta saved
Estimated Results are saved in ex.teddf.nonr.weight.result.dta.
.
```

5.2 Application of gtfpch

We first apply `gtfpch` to estimate the Malmquist–Luenberger productivity index (MLPI) to measure the green total-factor productivity growth of China's provinces. Regarding the results, TFPCH stores the values of MLPI; TECH and TECCH are the two decomposition terms of MLPI, describing technical efficiency change and technological change, respectively. Note that we implement the estimation based on the global technology benchmark by specifying the *global* option.

```
.
. egen id=group(Province)
. xtset id year
       panel variable:  id (strongly balanced)
        time variable:  year, 2013 to 2015
                delta:  1 unit
. gtfpch K L= Y: CO2, dmu(Province) global sav(ex.gtfpch.result,replace)
 The directional vector is (0 0 Y -CO2)

 Total Factor Productivity Change:Malmquist-Luenberger Productivity Index
    (Row: Row # in the original data; Pdwise: periodwise)

     | Row    Province   id      Pdwise    TFPCH     TECH    TECCH |
  1. |   2       Anhui    1   2013~2014   0.9943   0.9662   1.0290 |
  2. |   3       Anhui    1   2014~2015   0.9951   0.9645   1.0317 |
  3. |   5     Beijing    2   2013~2014   1.0328   1.0000   1.0328 |
  4. |   6     Beijing    2   2014~2015   1.0583   1.0000   1.0583 |
  5. |   8   Chongqing    3   2013~2014   1.0013   0.9883   1.0132 |
  6. |   9   Chongqing    3   2014~2015   1.0222   0.9659   1.0582 |
                                  ...
```

```
                                   ...
                                   ...
 55. | 83      Xinjiang  28  2013~2014  0.9897  0.9746  1.0154 |
 56. | 84      Xinjiang  28  2014~2015  0.9881  0.9727  1.0159 |
 57. | 86        Yunnan  29  2013~2014  1.0161  0.9889  1.0275 |
 58. | 87        Yunnan  29  2014~2015  1.0155  0.9849  1.0310 |
 59. | 89      Zhejiang  30  2013~2014  1.0143  0.9677  1.0481 |
 60. | 90      Zhejiang  30  2014~2015  1.0028  0.9695  1.0343 |
     +--------------------------------------------------------+

Note: missing value indicates infeasible problem.
(note: file ex.gtfpch.result.dta not found)
file ex.gtfpch.result.dta saved

Estimated Results are saved in ex.gtfpch.result.dta.

.
```

Alternatively, **gtfpch** can be employed to estimate the Luenberger productivity indicator. We present an example as follows.

```
.
. gtfpch K L= Y: CO2, dmu( Province ) nonr  global sav(ex.gtfpch.nonr.result,replace)

 The weight vector is (0 0 1 1)

 The directional vector is (0 0 Y -CO2)

 Total Factor Productivity Change:Luenberger Productivity Index (based on nonrial DDF)
    (Row: Row # in the original data; Pdwise: periodwise)
```

	Row	Province	id	Pdwise	TFPCH	TECH	TECCH
1.	2	Anhui	1	2013~2014	-0.0676	-0.2281	0.1605
2.	3	Anhui	1	2014~2015	0.0214	-0.0597	0.0811
3.	5	Beijing	2	2013~2014	0.0832	-0.0000	0.0832
4.	6	Beijing	2	2014~2015	0.1705	0.0000	0.1705
5.	8	Chongqing	3	2013~2014	0.0175	-0.0564	0.0738
6.	9	Chongqing	3	2014~2015	0.0178	-0.1079	0.1257
				...			
				...			
				...			
55.	83	Xinjiang	28	2013~2014	-0.2221	-0.3371	0.1150
56.	84	Xinjiang	28	2014~2015	-0.2232	-0.2931	0.0699
57.	86	Yunnan	29	2013~2014	0.0128	-0.1320	0.1448
58.	87	Yunnan	29	2014~2015	0.1378	0.0586	0.0792
59.	89	Zhejiang	30	2013~2014	0.0119	-0.0558	0.0677
60.	90	Zhejiang	30	2014~2015	-0.0092	-0.0996	0.0903

```
Note: missing value indicates infeasible problem.
(note: file ex.gtfpch.nonr.result.dta not found)
file ex.gtfpch.nonr.result.dta saved

Estimated Results are saved in ex.gtfpch.nonr.result.dta.

.
```

6 Conclusion

With the increasing demand for improving sustainability at the macro and micro levels, scholars and managers recognized that it is becoming more and more important to consider undesirable output in efficiency and productivity analysis.

Stata, as one of the leading packages for economic analysis, however, has not provided comprehensive tools to measure technical efficiency and total factor productivity change when considering undesirable outputs. Here, as an attempt to fill this gap, we introduced two new Stata commands that perform estimations for nonparametric frontier models with undesirable outputs.

`teddf` estimates directional distance function with undesirable outputs for technical efficiency measurement. Both radial Debreu-Farrell and non-radial Russell measures can be calculated, under different assumptions about the production technology, e.g., window, biennial, sequential, and global production technology. `gtfpch` measures total factor productivity change with undesirable outputs using Malmquist–Luenberger productivity index or Luenberger indicator. Two types of specifications of decomposing total factor productivity change were given. Some empirical examples have been presented to show the usage of the two commands.

Finally, it should be noted that the models we introduced are DEA-type estimators, which might be sensitive to sampling variation. Thus, the statistical inference in this context is critical. However, there are still many open issues both in theory and application.

7 Acknowledgments

Kerui Du thanks the financial support of the National Natural Science Foundation of China (72074184) and the Fundamental Research Funds for the Central Universities (20720201016). Ning Zhang thanks the financial support of the National Natural Science Foundation of China (72033005; 71822402). We are grateful to Stephen P. Jenkins and the anonymous reviewer for the helpful comments and suggestions which led to an improved version of this paper.

8 References

Badunenko, O., and P. Mozharovskyi. 2016. Nonparametric frontier analysis using Stata. *Stata Journal* 16(3): 550–589. URL http://www.stata-journal.com/sjpdf.html?article=st0444.

———. 2020. Statistical inference for the Russell measure of technical efficiency. *Journal of the Operational Research Society* 71(3): 517–527. URL https://doi.org/10.1080/01605682.2019.1599778.

Banker, R. D. 1993. Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation. *Management Science* 39(10): 1265–1273. URL http://www.jstor.org/stable/2632965.

Chambers, R. G. 2002. Exact nonradial input, output, and productivity measurement. *Economic Theory* 20(4): 751–765. URL http://link.springer.com/10.1007/s001990100231.

Chung, Y., R. Färe, and S. Grosskopf. 1997. Productivity and Undesirable Outputs: A Directional Distance Function Approach. *Journal of Environ-*

mental Management 51(3): 229–240. URL https://linkinghub.elsevier.com/retrieve/pii/S0301479797901468.

Debreu, G. 1951. The Coefficient of Resource Utilization. *Econometrica* 19(3): 273. URL https://www.jstor.org/stable/1906814?origin=crossref.

Färe, R., and S. Grosskopf. 2010. Directional distance functions and slacks-based measures of efficiency. *European Journal of Operational Research* 200(1): 320–322. URL https://linkinghub.elsevier.com/retrieve/pii/S0377221709000186.

Färe, R., S. Grosskopf, and C. A. K. Lovell. 1985. *The Measurement of Efficiency of Production*. Dordrecht: Springer Netherlands. URL http://link.springer.com/10.1007/978-94-015-7721-2.

Färe, R., S. Grosskopf, M. Norris, and Z. Zhang. 1994. Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries. *The American Economic Review* 84(1): 66–83. URL http://www.jstor.org/stable/2117971.

Farrell, M. J. 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)* 120(3): 253. URL https://www.jstor.org/stable/2343100?origin=crossref.

Ji, Y., and C. Lee. 2010. Data envelopment analysis. *Stata Journal* 10(2): 267–280. URL http://www.stata-journal.com/sjpdf.html?article=st0193.

Kneip, A., B. U. Park, and L. Simar. 1998. A Note on the Convergence of Nonparametric DEA Estimators for Production Efficiency Scores. *Econometric Theory* 14(6): 783–793. URL http://www.jstor.org/stable/3533091.

Kneip, A., L. Simar, and P. W. Wilson. 2008. ASYMPTOTICS AND CONSISTENT BOOTSTRAPS FOR DEA ESTIMATORS IN NONPARAMETRIC FRONTIER MODELS. *Econometric Theory* 24(6): 1663–1697.

Koopmans, T. C. 1951. An analysis of production as an efficient combination of activities. *Activity analysis of production and allocation* .

Lin, B., and K. Du. 2015. Modeling the dynamics of carbon emission performance in China: A parametric Malmquist index approach. *Energy Economics* 49: 550–557. URL https://linkinghub.elsevier.com/retrieve/pii/S0140988315001206.

Mahlberg, B., and B. K. Sahoo. 2011. Radial and non-radial decompositions of Luenberger productivity indicator with an illustrative application. *International Journal of Production Economics* 131(2): 721–726. URL https://linkinghub.elsevier.com/retrieve/pii/S092552731100079X.

Ray, S. C., and E. Desli. 1997. Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries: Comment. *The American Economic Review* 87(5): 1033–1039. URL http://www.jstor.org/stable/2951340.

Simar, L., A. Vanhems, and P. W. Wilson. 2012. Statistical inference for DEA estimators of directional distances. *European Journal of Operational Research* 220(3): 853–864. URL https://www.sciencedirect.com/science/article/pii/S0377221712001622.

Simar, L., and P. W. Wilson. 2019. Central limit theorems and inference for sources of productivity change measured by nonparametric Malmquist indices. *European Journal of Operational Research* 277(2): 756–769. URL https://www.sciencedirect.com/science/article/pii/S0377221719301882.

Simar, L., and P. W. Wilson. 1998. Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models. *Management Science* 44(1): 49–61.

———. 1999. Estimating and bootstrapping Malmquist indices. *European Journal of Operational Research* 115(3): 459–471. URL https://www.sciencedirect.com/science/article/pii/S0377221797004505.

Tauchmann, H. 2012. Partial frontier efficiency analysis. *Stata Journal* 12(3): 461–478. URL http://www.stata-journal.com/sjpdf.html?article=st0270.

Yan, Z., B. Zou, K. Du, and K. Li. 2020. Do renewable energy technology innovations promote China's green productivity growth? Fresh evidence from partially linear functional-coefficient models. *Energy Economics* 90: 104842. URL http://www.sciencedirect.com/science/article/pii/S0140988320301821.

Zhou, P., B. Ang, and H. Wang. 2012. Energy and CO2 emission performance in electricity generation: A non-radial directional distance function approach. *European Journal of Operational Research* 221(3): 625–635. URL https://linkinghub.elsevier.com/retrieve/pii/S0377221712003153.

About the authors

Daoping Wang is a PhD candidate at the School of Urban and Regional Science, Shanghai University of Finance and Economics. His primary research interests include economic dynamics, risk analysis, and sustainable development.

Kerui Du (corresponding author) is an associate professor at the School of Management, Xiamen University. His primary research interests include applied econometrics, energy and environmental economics.

Ning Zhang is a professor at the Institute of Blue and Green Development, Shandong University. His primary research interests include environmental economics and energy economics.