Centre for Efficiency and Productivity Analysis

Working Paper Series
No. WP01/2017

Estimating Total Factor Productivity Change When No Price or Value-Share Data are Available

C.J. O'Donnell

Date: January 2017

School of Economics
University of Queensland
St. Lucia, Qld. 4072
Australia

ISSN No. 1932 - 4398

Estimating Total Factor Productivity Change When No Price or Value-Share Data are Available

by

C.J. O'Donnell

Centre for Efficiency and Productivity Analysis, The University of Queensland

Email: c.odonnell@economics.uq.edu.au

23 January, 2017

1. Introduction

Total factor productivity (TFP) is a measure of total output divided by a measure of total input. A TFP index is a measure of total output change divided by a measure of total input change (i.e., an output index divided by an input index). It is common to compute output and input index numbers using formulas that combine quantity data with either price or value-share data. Examples include the well-known Fisher and Törnqvist indices. This paper explains how output and input index numbers (and therefore TFP index numbers) can be computed using quantity data only. It draws heavily on O'Donnell (2012, 2014, 2016).

This paper focuses on output and input indices that are "proper" in the sense that they satisfy the common notions, or axioms, listed in O'Donnell (2016). Included among these are a transitivity axiom and a proportionality axiom. The transitivity axiom says, for example, that if firm A produced the same outputs as firm B, and firm B produced the same outputs as firm C, then firm A produced the same outputs as firm C. The proportionality axiom says, for example, that if a firm used λ times as much of every input in period 2 as it did in period 1, then the index that measures the increase in input use must take the value λ. The class of proper output and input indices includes several types of additive, multiplicative and primal indices.

If there are no price or value-share data available, then the computation of output and input index numbers generally involves estimating output and input distance functions. This paper explains how to estimate distance functions using linear programming (LP), least squares (LS) and maximum likelihood (ML) methods. LP methods are widely used to estimate the parameters of piecewise frontier models (PFMs). LS methods are widely used to estimate the parameters of deterministic frontier models (DFMs). ML methods are widely used to estimate the parameters of stochastic frontier models (SFMs). The estimated parameters of PFMs and DFMs can be used to compute additive, multiplicative and primal index numbers. The estimated parameters of SFMs can generally only be used to compute additive and multiplicative index numbers.

The outline of the paper is as follows. Section 2 formally defines what is meant by a "technology". It then explains how sets of technologies can be represented by production possibilities sets and distance functions. Section 3 defines several additive, multiplicative and primal indices. It also explains how TFP index numbers can be computed using estimates of shadow prices and shadow value shares. Section 4 explains how shadow prices and value shares can be estimated using common LP methods. It then uses a numerical example to illustrate some of the properties of proper TFP indices. Sections 5 explains how LS estimators can be used to estimate the parameters of distance functions. It then shows how the estimated parameters of either an output distance function or an input distance function can be used to compute additive, multiplicative and primal TFP index numbers. Section 6 explains how ML estimators can be used to estimate the parameters of stochastic frontier models. In this case, the estimated parameters can only be used to compute additive and multiplicative index numbers. Finally, Section 7 explains how LP and LS methods can be used to compute several other measures of TFP change. These other measures include generalised Malmquist TFP indices, Hicks-Moorsteen TFP indices, and residual measures of productivity change.

2. Technologies

In O'Donnell (2016), a technology is defined as "a technique, method or system for transforming inputs into outputs ... For most practical intents and purposes, it is convenient to think of a technology as a book of instructions, or recipe" (p. 328). In this paper,[1] the set of technologies that exist in a given period is referred to as a technology set. If we think of a technology as a book of instructions, then we can follow O'Donnell

[1] In O'Donnell (2016), technology sets are referred to as "metatechnologies".

(2016) and think of a technology set as a library. Technology sets can be represented by production possibilities sets.

2.1. *Production Possibilities Sets*

A production possibilities set is a set containing all input-output combinations that are physically possible. The building blocks for this paper are *period-and-environment-specific* production possibilities sets. A period-and-environment-specific production possibilities set is a set containing all input-output combinations that are physically possible *in a given period in a given production environment*. For a precise definition, let $z_{it} = (z_{1it}, \ldots, z_{Jit})'$ denote a vector of environmental variables characterising the production environment of firm i in period t. Here, the term "environmental variables" refers to "variables that are physically involved in the production process but never chosen by firms (e.g., rainfall in farming, or the highway network in trucking)" (O'Donnell, 2016, p.329). The set of input-output combinations that are possible using the period-t technology set in an environment characterised by z_{it} is

$$T^t(z_{it}) = \{(x, q) : \text{inputs } x \text{ can produce outputs } q \text{ in period } t \text{ in environment } z_{it}\}. \quad (1)$$

If there is no technical change, then all references to period t can be deleted. If there are no environmental variables involved in the production process (i.e., if there is no environmental change), then all references to z_{it} can be deleted. If there is no technical or environmental change, then $T^t(z_{it})$ is equal to the "technology set" defined by Färe and Primont (1995, p. 8), the "production set" defined by Fried et al. (2008, Eq. 1.1), and the "graph" defined by Shephard (1970, p. 181), Färe, Grosskopf and Lovell (1994, Eq. 2.1.5) and Kumbhakar and Lovell (2000, p. 18).

Economists make assumptions about technology sets by way of assumptions about what they can (and cannot) produce. For example, it is common to make the following assumptions: (a) it is possible to produce zero output (i.e., "inactivity is possible"); (b) there is a limit to what can be produced using a finite amount of inputs (i.e., "output sets are bounded"); (c) a positive amount of at least one input is needed in order to produce a positive amount of any output (i.e., "there is no free lunch"); (d) if given inputs can be used to produce a particular output vector, then they can also be used to produce a scalar contraction of that output vector (i.e., "outputs are weakly disposable"); (e) if given outputs can be produced using a particular input vector, then they can also be produced using a scalar magnification of that input vector (i.e.,“ inputs are weakly disposable”); and (f) output and input sets contain all the points on their boundaries (i.e., they are “closed”). Stronger assumptions that are made from time to time include the following: (g) if given inputs can be used to produce particular outputs, then they can also be used to produce fewer outputs (i.e., “outputs are strongly disposable”); and (h) if given outputs can be produced using particular inputs, then they can also be produced using more inputs (i.e., “inputs are strongly disposable”). If outputs (resp. inputs) are strongly disposable, then they are also weakly disposable. If outputs are weakly disposable and output sets are bounded, then production possibilities sets can be represented by output distance functions. If inputs are weakly disposable, then they can be represented by input distance functions.

2.2. Output Distance Functions

An output distance function (ODF) gives the reciprocal of the largest factor by which a firm can scale up its output vector when its inputs are predetermined. ODFs can be traced back to Shephard (1970, p. 207). A *period-and-environment-specific* ODF gives the reciprocal of the largest factor by which a firm can scale up its output vector *in a given period in a given production environment* when its inputs are predetermined. For a mathematical definition, let $x_{it} = (x_{1it}, \ldots, x_{Mit})' \geq 0$ and $q_{it} = (q_{1it}, \ldots, q_{Nit})' \geq 0$ denote the input and output vectors of firm i in period t. In this paper, the period-and-environment-specific ODF is defined as

$$D_O^t(x_{it}, q_{it}, z_{it}) = \inf\{\rho > 0 : (x_{it}, q_{it}/\rho) \in T^t(z_{it})\}. \tag{2}$$

If there is no environmental change, then this function is equal to the ODF defined by Balk (1998, Eq. 2.6). If output sets are closed and there is no technical or environmental change, then it is equal to the distance function defined by Shephard (1970, p. 207), Färe and Primont (1995, Eq. 2.1.6) and Kumbhakar and Lovell (2000, p. 30).

If it exists, then the ODF defined by (2) is nonnegative and linearly homogeneous in outputs. If outputs are strongly disposable, then it is also nondecreasing in outputs. If outputs are strongly disposable and the ODF is continuously differentiable, then the n-th normalised shadow output price and shadow revenue share are $p_n^t(x_{it}, q_{it}, z_{it}) = \partial D_O^t(x_{it}, q_{it}, z_{it})/\partial q_{nit} \geq 0$ and $r_n^t(x_{it}, q_{it}, z_{it}) = \partial \ln D_O^t(x_{it}, q_{it}, z_{it})/\partial \ln q_{nit} \geq 0$. Shadow output prices are the prices that would lead a price-taking revenue-maximising firm to choose its observed outputs. Normalised shadow output prices are shadow output prices divided by the maximum revenue the firm can earn using its inputs. As we shall see, shadow output prices and revenue shares play an important role in the computation of output index numbers.

2.3. Input Distance Functions

An input distance function (IDF) gives the reciprocal of the smallest fraction of inputs that a firm needs to produce its outputs. IDFs can be traced back to Shephard (1953, p. 5; 1970, p. 206). A *period-and-environment-specific* IDF gives the reciprocal of the smallest fraction of inputs that a firm needs to produce its outputs *in a given period in a given production environment*. In this paper, the period-and-environment-specific IDF is defined as

$$D_I^t(x_{it}, q_{it}, z_{it}) = \sup\{\theta > 0 : (x_{it}/\theta, q_{it}) \in T^t(z_{it})\}. \tag{3}$$

If the production frontier exhibits constant returns to scale (CRS), then $D_I^t(x_{it}, q_{it}, z_{it}) = 1/D_O^t(x_{it}, q_{it}, z_{it})$. If there is no environmental change, then the IDF defined by (3) is equal to the IDF defined by Balk (1998, Eq. 2.4). If input sets are closed and there is no technical or environmental change, then it is equal to the IDF defined by Shephard (1970, p. 206) and Kumbhakar and Lovell (2000, p. 28).

If it exists, then the IDF defined by (3) is nonnegative and linearly homogeneous in inputs. If inputs are strongly disposable, then it is also nondecreasing in inputs. If inputs are strongly disposable and the IDF is continuously differentiable, then the m-th normalised shadow input price is $w_m^t(x_{it}, q_{it}, z_{it}) = \partial D_I^t(x_{it}, q_{it}, z_{it})/\partial x_{mit} \geq 0$ and the m-th shadow cost share is $s_m^t(x_{it}, q_{it}, z_{it}) = \partial \ln D_I^t(x_{it}, q_{it}, z_{it})/\partial \ln x_{mit} \geq 0$. Shadow input prices are the prices that would lead a price-taking cost-minimising firm to choose its observed inputs. Normalised shadow input prices are shadow input prices divided by the minimum cost of producing the firm's outputs. Shadow input prices and cost shares play an important role in the computation of *input* index numbers.

2.4. Example

For a simple example that is compatible with the numerical illustrations presented later in the paper, suppose there is only one environmental variable. Also suppose that the set of input-output combinations that are possible in period t in an environment characterised by z_{it} is

$$T^t(z_{it}) = \left\{ (x,q) : \left(\sum_{n=1}^{N} \gamma_n q_n \right) \leq A(t)\, z_{it}^{\delta} \prod_{m=1}^{M} x_m^{\beta_m} \right\} \tag{4}$$

where $A(t) \geq A(t-1) > 0$, $\beta = (\beta_1, \ldots, \beta_M)' \geq 0$, $\gamma = (\gamma_1, \ldots, \gamma_N)' \geq 0$ and $\sum_n \gamma_n = 1$. The constraint $A(t) \geq A(t-1)$ implies there is no technical regress. The constraints $\beta \geq 0$ and $\gamma \geq 0$ imply that inputs and outputs are strongly disposable. The associated ODF and IDF are

$$D_O^t(x_{it}, q_{it}, z_{it}) = \left(\sum_{n=1}^{N} \gamma_n q_{nit} \right) \left(A(t)\, z_{it}^{\delta} \prod_{m=1}^{M} x_{mit}^{\beta_m} \right)^{-1} \tag{5}$$

$$\text{and} \quad D_I^t(x_{it}, q_{it}, z_{it}) = \left(B(t)\, z_{it}^{\kappa} \prod_{m=1}^{M} x_{mit}^{\lambda_m} \right) \left(\sum_{n=1}^{N} \gamma_n q_{nit} \right)^{-1/\eta} \tag{6}$$

where $\eta = \sum_m \beta_m > 0$, $B(t) = A(t)^{1/\eta} > 0$, $\lambda = (\lambda_1, \ldots, \lambda_M)' = \beta/\eta \geq 0$, and $\kappa = \delta/\eta$. If $\eta = 1$, then, and only then, the production frontier exhibits CRS. The n-th normalised shadow output price and shadow revenue share are $p_n^t(x_{it}, q_{it}, z_{it}) = \gamma_n D_O^t(x_{it}, q_{it}, z_{it})/\gamma' q_{it}$ and $r_n(q_{it}) = \gamma_n q_{nit}/\gamma' q_{it}$. Here, the notation reflects the fact that shadow revenue shares do not vary over time and do not depend on inputs or environmental variables. The m-th normalised shadow input price and shadow cost share are $w_m^t(x_{it}, q_{it}, z_{it}) = \lambda_m D_I^t(x_{it}, q_{it}, z_{it})/x_{mit}$ and $s_m = \lambda_m$. Here, the notation reflects the fact that shadow cost shares do not vary over firms or time periods.

3. TFP Indices

An index that compares the outputs of firm i in period t with the outputs of firm k in period s is any variable of the form $QI(q_{ks}, q_{it}) = Q(q_{it})/Q(q_{ks})$ where $Q(.)$ is a nonnegative, nondecreasing, linearly-homogeneous, scalar-valued aggregator function. Output indices that are constructed in this way satisfy the monotonicity, homogeneity, identity, proportionality, time-space reversal, transitivity and circularity axioms listed in O'Donnell (2016, p. 332). O'Donnell (2016) refers to an index as being "proper" if and only if *all* of these axioms are satisfied.

On the input side, an index that compares the inputs of firm i in period t with the inputs of firm k in period s is any variable of the form $XI(x_{ks}, x_{it}) = X(x_{it})/X(x_{ks})$ where $X(.)$ is a nonnegative, nondecreasing, linearly-homogeneous, scalar-valued aggregator function. Again, input indices that are constructed in this way are "proper" in the sense that they satisfy the input axioms listed in O'Donnell (2016, p. 332).

Finally, a "proper" index that compares the TFP of firm i in period t with the TFP of firm k in period s is any variable of the form $TFPI(x_{ks}, q_{ks}, x_{it}, q_{it}) = QI(q_{ks}, q_{it})/XI(x_{ks}, x_{it})$ where $QI(q_{ks}, q_{it})$ and $XI(x_{ks}, x_{it})$ are proper output and input indices. Proper TFP indices satisfy their own set of axioms, including the following:

T1 $q_{rl} \geq q_{it}$ and $x_{rl} \leq x_{it} \Rightarrow TFPI(x_{ks}, q_{ks}, x_{rl}, q_{rl}) \geq TFPI(x_{ks}, q_{ks}, x_{it}, q_{it})$ (weak monotonicity);

T2 $q_{rl} = \lambda q_{it}$ and $x_{rl} = \delta x_{it} \Rightarrow TFPI(x_{ks}, q_{ks}, x_{rl}, q_{rl}) = (\lambda/\delta) TFPI(x_{ks}, q_{ks}, x_{it}, q_{it})$ for $\lambda > 0$ and $\delta > 0$ (homogeneity);

T3 $q_{ks} = q_{it}$ and $x_{ks} = x_{it} \Rightarrow TFPI(x_{ks}, q_{ks}, x_{it}, q_{it}) = 1$ (identity);

T4 $q_{hj} = \lambda q_{ks}$, $q_{rl} = \lambda q_{it}$, $x_{hj} = \delta x_{ks}$ and $x_{rl} = \delta x_{it} \Rightarrow$ $TFPI(x_{hj}, q_{hj}, x_{rl}, q_{rl}) = TFPI(x_{ks}, q_{ks}, x_{it}, q_{it})$ for $\lambda > 0$ and $\delta > 0$;

T5 $q_{it} = \lambda q_{ks}$ and $x_{it} = \delta x_{ks} \Rightarrow TFPI(x_{ks}, q_{ks}, x_{it}, q_{it}) = \lambda/\delta$ for $\lambda > 0$ and $\delta > 0$ (proportionality);

T6 $TFPI(x_{ks}, q_{ks}, x_{it}, q_{it}) = 1/TFPI(x_{it}, q_{it}, x_{ks}, q_{ks})$ (time-space reversal);

T7 $TFPI(x_{ks}, q_{ks}, x_{it}, q_{it}) = TFPI(x_{ks}, q_{ks}, x_{rl}, q_{rl}) TFPI(x_{rl}, q_{rl}, x_{it}, q_{it})$ (transitivity); and

T8 $TFPI(x_{ks}, q_{ks}, x_{rl}, q_{rl}) TFPI(x_{rl}, q_{rl}, x_{it}, q_{it}) TFPI(x_{it}, q_{it}, x_{ks}, q_{ks}) = 1$ (circularity).

The interpretation of these axioms is straightforward. The identity axiom, for example, says that if the inputs and outputs of firm i in period t are exactly the same as the inputs and outputs of firm k in period s, then the index that compares the TFP of the two firms must take the value 1. The transitivity axiom says, for example, that if firm i in period t is twice as productive as firm r in period l, and if firm r in period l is twice as productive as firm k in period s, then all direct and indirect comparisons must say that firm i in period t is four times as productive as firm k in period s. The class of proper TFP indices includes various additive, multiplicative and primal indices.

3.1. *Additive Indices*

The use of the term "additive" in an index number context can be traced back at least as far as Aczel and Eichhorn (1974). Additive TFP indices are constructed using aggregator functions of the form $Q(q_{it}) \propto \sum_n a_n q_{nit}$ and $X(x_{it}) \propto \sum_m b_m x_{mit}$ where $a_n \geq 0$ and $b_m \geq 0$. The associated index that compares the TFP of firm i in period t with the TFP of firm k in period s is

$$TFPI^A(x_{ks}, q_{ks}, x_{it}, q_{it}) = \left(\frac{\sum_n a_n q_{nit}}{\sum_n a_n q_{nks}} \right) \left(\frac{\sum_n b_m x_{mks}}{\sum_m b_m x_{mit}} \right). \tag{7}$$

Any nonnegative measures of relative value can be used as weights. If average market prices are used as weights, then $TFPI^A(x_{ks}, q_{ks}, x_{it}, q_{it})$ takes the form of the Lowe TFP index defined by O'Donnell (2012, p.877). O'Donnell (2012) uses the term "Lowe" because the component output and input indices in his TFP index have a structure that is commonly attributed to Lowe (1823); see, for example, Hill (2008) and Balk (2008, p. 68). If average market prices are unavailable and average normalised shadow prices are used as weights, then,

$$TFPI^A(x_{ks}, q_{ks}, x_{it}, q_{it}) = \left(\frac{\sum_n \bar{p}_n q_{nit}}{\sum_n \bar{p}_n q_{nks}} \right) \left(\frac{\sum_n \bar{w}_m x_{mks}}{\sum_m \bar{w}_m x_{mit}} \right) \tag{8}$$

where $\bar{p}_n \propto \sum_i \sum_t p_n^t(x_{it}, q_{it}, z_{it})$ and $\bar{w}_m \propto \sum_i \sum_t w_m^t(x_{it}, q_{it}, z_{it})$.

3.2. *Multiplicative Indices*

The use of the term "multiplicative" in an index number context can be traced back at least as far as Coelli et al. (2005, p. 131). Multiplicative TFP indices are constructed using aggregator functions of the form $Q(q_{it}) \propto \prod_{n=1}^N q_{nit}^{a_n}$ and $X(x_{it}) \propto \prod_{m=1}^M x_{mit}^{b_m}$ where $a_n \geq 0$, $b_m \geq 0$ and $\sum_n a_n = \sum_m b_m = 1$. The associated index that compares the TFP of firm i in period t with the TFP of firm k in period s is

$$TFPI^M(x_{ks}, q_{ks}, x_{it}, q_{it}) = \prod_{n=1}^N \left(\frac{q_{nit}}{q_{nks}} \right)^{a_n} \prod_{m=1}^M \left(\frac{x_{mks}}{x_{mit}} \right)^{b_m}. \tag{9}$$

Any nonnegative measures of relative value *that sum to one* can be used as weights. Let I_t denote the number of firms in period t. If average shadow revenue and cost shares are used as weights, then

$$TFPI^M(x_{ks}, q_{ks}, x_{it}, q_{it}) = \prod_{n=1}^{N} \left(\frac{q_{nit}}{q_{nks}} \right)^{\bar{r}_n} \prod_{m=1}^{M} \left(\frac{x_{mks}}{x_{mit}} \right)^{\bar{s}_m} \tag{10}$$

where $\bar{r}_n = \sum_i \sum_t r_n^t(x_{it}, q_{it}, z_{it}) / (\sum_t I_t)$ and $\bar{s}_m = \sum_i \sum_t s_m^t(x_{it}, q_{it}, z_{it}) / (\sum_t I_t)$. If average observed value shares are equal to average shadow value shares, then, and only then, $TFPI^M(x_{ks}, q_{ks}, x_{it}, q_{it})$ is equivalent to the geometric Young TFP index defined by O'Donnell (2016, Eq. 5). The term "geometric Young" can be traced back to a producer price index described by IMF (2004, p.10).

3.3. Primal Indices

The use of the term "primal" in an index number context can be traced back at least as far as Balk (1998, p.100).[2] Primal TFP indices are constructed using output and input distance functions. To be more precise, if outputs and inputs are strongly disposable, then suitable aggregator functions are $Q(q_{it}) \propto D_O^{\bar{t}}(\bar{x}, q_{it}, \bar{z})$ and $X(x_{it}) \propto D_I^{\bar{t}}(x_{it}, \bar{q}, \bar{z})$ where $\bar{t}$ is a representative time period and $\bar{x}$, $\bar{q}$ and $\bar{z}$ are representative vectors of inputs, outputs and environmental variables. The associated (primal) index that compares the TFP of firm i in period t with the TFP of firm k in period s is

$$TFPI^P(x_{ks}, q_{ks}, x_{it}, q_{it}) = \frac{D_O^{\bar{t}}(\bar{x}, q_{it}, \bar{z})}{D_O^{\bar{t}}(\bar{x}, q_{ks}, \bar{z})} \frac{D_I^{\bar{t}}(x_{ks}, \bar{q}, \bar{z})}{D_I^{\bar{t}}(x_{it}, \bar{q}, \bar{z})}. \tag{11}$$

This index appears in O'Donnell (2016, p.332). If there is no technical or environmental change, then the two ratios in (11) are equal to the output and input quantity indices of Färe and Primont (1995, pp. 36, 38). Other special cases are discussed in O'Donnell (2016).

[2]The use of the term "primal" in this paper is also consistent with the way the term is use in economics. In economics, the term is usually used to describe production functions, input requirement functions and distance functions. The term "dual" is usually reserved for cost, revenue and profit functions. See, for example, Beavis and Dobbs (1990, p. 99).

3.4. Example

If the production possibilities set is given by (4), then the ODF and IDF are given by (5) and (6). In this case, the additive index defined by (8) can be written as

$$TFPI^{A}(x_{ks},q_{ks},x_{it},q_{it}) = \left(\frac{\sum_n \gamma_n q_{nit}}{\sum_n \gamma_n q_{nks}}\right)\left(\frac{\sum_m \bar{w}_m x_{mks}}{\sum_m \bar{w}_m x_{mit}}\right) \tag{12}$$

where $\bar{w}_m \propto \sum_i \sum_t (\lambda_m / x_{mit})$. The multiplicative index defined by (10) can be written as

$$TFPI^{M}(x_{ks},q_{ks},x_{it},q_{it}) = \prod_{n=1}^{N}\left(\frac{q_{nit}}{q_{nks}}\right)^{\bar{r}_n}\prod_{m=1}^{M}\left(\frac{x_{mks}}{x_{mit}}\right)^{\lambda_m} \tag{13}$$

where $\bar{r}_n = \sum_i \sum_t (\gamma_n q_{nit}/\gamma' q_{it})/(\sum_t I_t)$. Finally, the primal index defined by (11) is

$$TFPI^{P}(x_{ks},q_{ks},x_{it},q_{it}) = \left(\frac{\sum_n \gamma_n q_{nit}}{\sum_n \gamma_n q_{nks}}\right)\prod_{m=1}^{M}\left(\frac{x_{mks}}{x_{mit}}\right)^{\lambda_m}. \tag{14}$$

Observe that this last index does not depend on $\bar{t}$, $\bar{x}$, $\bar{q}$ or $\bar{z}$. This is due to the fact that output and input sets are homothetic and technical change is Hicks neutral.

4. Linear Programming (LP) Estimators

If no price or value share data are available, then the computation of TFP index numbers generally involves the estimation of distance functions. In the productivity literature, is common to assume that

LP1: all relevant quantities, prices and environmental variables are observed and measured without error;

LP2: production frontiers are locally linear;

LP3: outputs and inputs are strongly disposable; and

LP4: production possibilities sets are convex.

If these assumptions are true, then estimates of the parameters of output and input distance functions can be obtained by solving linear programs. The associated models are known as data envelopment analysis (DEA) models. The idea behind DEA can be traced back to Farrell (1957) and Farrell and Fieldhouse (1962). The formulation of DEA estimation problems as linear programs is largely due to Charnes et al. (1978) and Banker et al. (1984).

4.1. Output-Oriented DEA Models

If LP1 and LP2 are true, then the ODF takes the following form:

$$D_O^{it}(x_{it}, q_{it}, z_{it}) = (\gamma_{it}' q_{it})/(\alpha_{it} + \delta_{it}' z_{it} + \beta_{it}' x_{it}). \tag{15}$$

Here, the i and t superscripts indicate that the parameters of the ODF vary with i and t. The term "locally linear" derives from the fact that if $D_O^{it}(x_{it}, q_{it}, z_{it}) = 1$ (i.e., the firm operates on the frontier), then the relationship between the inputs, outputs and environmental variables can be written as $\gamma_{it}' q_{it} = \alpha_{it} + \delta_{it}' z_{it} + \beta_{it}' x_{it}$ (i.e., a linear function with observation-varying parameters).

The unknown parameters in (15) are not identified. In practice, it is common to identify the parameters by setting $\gamma_{it}' q_{it} = 1$. LP estimation of the parameters then involves minimising $1/D_O^{it}(x_{it}, q_{it}, z_{it})$ subject to constraints that ensure that LP3 and LP4 are also satisfied. The estimation problem is

$$\begin{aligned} \min_{\alpha_{it}, \delta_{it}, \beta_{it}, \gamma_{it}} \Big\{ & \alpha_{it} + \delta_{it}' z_{it} + \beta_{it}' x_{it} : \gamma_{it} \geq 0,\ \beta_{it} \geq 0,\ \gamma_{it}' q_{it} = 1, \\ & \alpha_{it} + \delta_{it}' z_{hr} + \beta_{it}' x_{hr} - \gamma_{it}' q_{hr} \geq 0 \text{ for all } h \leq I \text{ and } r \leq t \Big\} \end{aligned} \tag{16}$$

where I is the number of firms in the dataset. This is a standard linear program. The constraints $\gamma_{it} \geq 0$ and $\beta_{it} \geq 0$ ensure that LP3 is satisfied. The last set of constraints ensures that LP4 is satisfied. If the data are cross section data, then all references to period t can be suppressed. If the data are cross section data and $\alpha_{it} + \delta_{it}' z_{it} = 0$, then (16) collapses to an LP that can be traced back to Charnes et al. (1978, Eq. 4).

LP (16) is sometimes referred to as a "primal" LP. Every primal LP has a "dual" form with the property that if the primal and dual forms both have feasible solutions, then the optimised values of the two objective functions are equal. The dual form of

(16) is

$$\max_{\mu,\lambda}\Big\{\mu:\ \mu q_{it}\leq\sum_{h=1}^{I}\sum_{r=1}^{t}\lambda_{hr}q_{hr},\ \sum_{h=1}^{I}\sum_{r=1}^{t}\lambda_{hr}x_{hr}\leq x_{it},$$
$$\sum_{h=1}^{I}\sum_{r=1}^{t}\lambda_{hr}z_{hr}=z_{it},\ \sum_{h=1}^{I}\sum_{r=1}^{t}\lambda_{hr}=1,\ \lambda_{hr}\geq 0 \text{ for all } h \text{ and } r\Big\}. \qquad (17)$$

Again, this is a standard LP. If there is no environmental change, then the environmental variables and their coefficients can be deleted. If the data are cross section data and there is no environmental change, then (17) collapses to an LP that can be found in, for example, Färe, Grosskopf and Lovell (1994, p. 103).

4.2. Input-Oriented DEA Models

If LP1 and LP2 are true, then the IDF is

$$D_I^{it}(x_{it},q_{it},z_{it})=(\theta_{it}'x_{it})/(\xi_{it}'q_{it}-\kappa_{it}'z_{it}-\phi_{it}). \qquad (18)$$

Again, the unknown parameters are not identified. In this case, it is common to identify the parameters by setting $\theta_{it}'x_{it}=1$. LP estimation of the parameters then involves maximising $1/D_I^{it}(x_{it},q_{it},z_{it})$ subject to constraints that ensure LP3 and LP4 are also satisfied. The estimation problem is

$$\max_{\xi_{it},\kappa_{it},\phi_{it},\theta_{it}}\Big\{\xi_{it}'q_{it}-\kappa_{it}'z_{it}-\phi_{it}:\xi_{it}\geq 0,\ \theta_{it}\geq 0,\ \theta_{it}'x_{it}=1,$$
$$\xi_{it}'q_{hr}-\kappa_{it}'z_{hr}-\phi_{it}-\theta_{it}'x_{hr}\leq 0 \text{ for all } h\leq I \text{ and } r\leq t\Big\}. \qquad (19)$$

Again, this is a standard linear program. Again, the constraints $\gamma_{it}\geq 0$ and $\beta_{it}\geq 0$ ensure that LP3 is satisfied. Again, the last set of constraints ensures that LP4 is satisfied. If the data are cross section data and there there is no environmental change, then the problem collapses to the dual LP of Banker et al. (1984, Eq. 20). Finally, the dual form of (19) is

$$\min_{\mu,\lambda}\Big\{\mu:\ \sum_{h=1}^{I}\sum_{r=1}^{t}\lambda_{hr}q_{hr}\geq q_{it},\ \mu x_{it}\geq\sum_{h=1}^{I}\sum_{r=1}^{t}\lambda_{hr}x_{hr},$$
$$\sum_{h=1}^{I}\sum_{r=1}^{t}\lambda_{hr}z_{hr}=z_{it},\ \sum_{h=1}^{I}\sum_{r=1}^{t}\lambda_{hr}=1,\ \lambda_{hr}\geq 0 \text{ for all } h \text{ and } r\Big\}. \qquad (20)$$

If the data are cross section data and there is no environmental change, then this problem collapses to the LP of Banker et al. (1984, Eq. 19).

4.3. TFP Indices

LP methods can be used to compute additive and multiplicative TFP index numbers. The computation of these index numbers is based on the result that if LP1–LP3 are true, then, by Euler's homogeneous function theorem,

$$D_O^{it}(x_{it}, q_{it}, z_{it}) = \sum_{n=1}^{N} p_n^{it}(x_{it}, q_{it}, z_{it}) q_{nit} \tag{21}$$

$$\text{and} \quad D_I^{it}(x_{it}, q_{it}, z_{it}) = \sum_{m=1}^{M} w_m^{it}(x_{it}, q_{it}, z_{it}) x_{mit} \tag{22}$$

where $p_n^{it}(x_{it}, q_{it}, z_{it}) = \gamma_{nit}/(\alpha_{it} + \delta_{it}' z_{it} + \beta_{it}' x_{it})$ is the n-th normalised shadow output price and $w_m^{it}(x_{it}, q_{it}, z_{it}) = \theta_{mit}/(\xi_{it}' q_{it} - \kappa_{it}' z_{it} - \phi_{it})$ is the m-th normalised shadow input price. Estimates of these shadow prices and shares can be obtained by solving problems (16) and (19). Additive TFP index numbers can then be computed using aggregator functions of the form $Q(q_{it}) \propto \hat{p}' q_{it}$ and $X(x_{it}) \propto \hat{w}' x_{it}$ where $\hat{p}$ is a vector of average estimated normalised shadow output prices and $\hat{w}$ is a vector of average estimated normalised shadow input prices. If average estimated normalised shadow prices are proportional to average market prices, then, and only then, the associated TFP index is equivalent to a Lowe TFP index. Similarly, multiplicative TFP index numbers can be computed using aggregator functions of the form $Q(q_{it}) \propto \prod_n q_{nit}^{\hat{r}_n}$ and $X(x_{it}) \propto \prod_m x_{mit}^{\hat{s}_m}$ where $\hat{r}_n$ is the n-th average estimated shadow revenue share and $\hat{s}_m$ is the m-th average estimated shadow cost share. If average estimated shadow value shares are equal to average observed value shares, then, and only then, the associated TFP index is equivalent to a geometric Young index.

LP methods can also be used to compute primal TFP index numbers. If LP1–LP3 are true, then $D_O^{\bar{t}}(\bar{x}, \bar{q}, \bar{z}) = \bar{p}' \bar{q}$ and $D_I^{\bar{t}}(\bar{x}, \bar{q}, \bar{z}) = \bar{w}' \bar{x}$ where $\bar{p} = \bar{\gamma}/(\bar{\alpha} + \bar{\delta}' \bar{z} + \bar{\beta}' \bar{x})$ is a vector of representative normalised shadow output prices and $\bar{w} = \bar{\theta}/(\bar{\xi}' \bar{q} - \bar{\kappa}' \bar{z} - \bar{\phi})$ is a vector of representative normalised shadow input prices. Estimates of $\bar{\alpha}$, $\bar{\delta}$, $\bar{\beta}$ and $\bar{\gamma}$

can be obtained as the solution to the following variant of problem (16):

$$\min_{\alpha,\delta,\beta,\gamma} \Big\{ \alpha + \delta'\bar{z} + \beta'\bar{x} : \gamma \geq 0,\ \beta \geq 0,\ \gamma'\bar{q} = 1,$$
$$\alpha + \delta' z_{hr} + \beta' x_{hr} - \gamma' q_{hr} \geq 0 \text{ for all } h \leq I \text{ and } r \leq \bar{t} \Big\}. \tag{23}$$

Similarly, estimates of $\bar{\xi}$, $\bar{\kappa}$, $\bar{\phi}$ and $\bar{\theta}$ can be obtained as the solution to the following variant of problem (19):

$$\max_{\xi,\kappa,\phi,\theta} \Big\{ \xi'\bar{q} - \kappa'\bar{z} - \phi : \xi \geq 0,\ \theta \geq 0,\ \theta'\bar{x} = 1,$$
$$\xi' q_{hr} - \kappa' z_{hr} - \phi - \theta' x_{hr} \leq 0 \text{ for all } h \leq I \text{ and } r \leq \bar{t} \Big\}. \tag{24}$$

An additive TFP index can then be computed using aggregator functions of the form $Q(q_{it}) \propto D_O^{\bar{t}}(\bar{x}, q_{it}, \bar{z}) = \bar{p}' q_{it}$ and $X(x_{it}) \propto D_I^{\bar{t}}(x_{it}, \bar{q}, \bar{z}) = \bar{w}' x_{it}$. The associated TFP index can be viewed as an estimate of a primal index.

4.4. Example

For a numerical illustration, Table 1 reports simulated data on one environmental variable and the quantities of $N = 2$ outputs and $M = 2$ inputs for $I = 4$ firms for $T = 5$ periods. Selected estimates of normalised shadow prices and value shares are reported in Table 2; these estimates were obtained by solving LPs (16) and (19). Associated TFP index numbers are reported in Table 3. The additive (resp. multiplicative) index numbers reported in column A (resp. M) were computed using the average estimated normalised shadow prices (resp. value shares) reported in the last line of Table 2. The primal index numbers reported in column P were computed using $\bar{p} = (0.675, 0.022)'$ and $\bar{w} = (0.099, 0.915)'$; these estimates were obtained by solving LPs (23) and (24).

The index numbers reported in Table 3 compare the TFP of each firm in each period with the TFP of firm 1 in period 1. In any given row, any differences between the additive and multiplicative index numbers are due to the use of different functional forms and weights. Any differences between the additive and primal index numbers are due to differences between average estimated normalised shadow prices (i.e., $\hat{p}$ and $\hat{w}$) and estimated representative shadow prices (i.e., $\bar{p}$ and $\bar{w}$). The index numbers in any given column provide a coherent picture of TFP change. For example, in any given column, the index number in row 17 correctly indicates that the TFP of firm 1 in period 5 was double what it had been in period 1 (in period 5, it used the same inputs to produce twice as much output as it had in period 1), the index numbers in rows 2 and 18 correctly indicate that firm 2 was equally productive in both periods (it used the same inputs in both periods to produce the same outputs), the index numbers in rows 3 and 19 correctly indicate that firm 3 was also equally productive in both periods (in period 5, this firm doubled its inputs in order to double the outputs it had produced in period 1), and the index numbers in rows 4 and 20 correctly indicate that firm 4 was equally productive in both periods (again, it used the same inputs in both periods to produce the same outputs).

Table 1: Simulated Data

Row	Firm	Period	q_1	q_2	x_1	x_2	z
1	1	1	1	1	1	1	1.43
2	2	1	0.41	0.5	0.22	0.47	0.99
3	3	1	0.13	0.11	0.07	0.18	0.85
4	4	1	0.3	1	0.21	1.01	1.03
5	1	2	2.09	2.09	2.09	2.09	1.72
6	2	2	1.49	1.58	1.3	1.55	1.38
7	3	2	0.8	0.78	0.74	0.85	1.18
8	4	2	0.39	1.09	0.3	1.1	1.07
9	1	3	1.75	1.75	1.75	1.75	1.64
10	2	3	1.99	2.08	1.8	2.05	1.51
11	3	3	0.95	0.93	0.89	1	1.23
12	4	3	2.91	3.61	2.82	3.62	1.76
13	1	4	1.01	1.01	1.01	1.01	1.43
14	2	4	2.02	2.11	1.83	2.08	1.52
15	3	4	0.48	0.46	0.42	0.53	1.04
16	4	4	1.48	2.18	1.39	2.19	1.42
17	1	5	2	2	1	1	0.97
18	2	5	0.41	0.5	0.22	0.47	1.25
19	3	5	0.26	0.22	0.14	0.36	0.84
20	4	5	0.3	1	0.21	1.01	1.38

Table 2: LP Estimates of Normalised Shadow Prices and Value Shares

Row	p_1	p_2	w_1	w_2	r_1	r_2	s_1	s_2
1	1	0	0.240	0.760	1	0	0.240	0.760
2	2.439	0	0.585	1.854	1	0	0.129	0.871
3	7.692	0	0	5.556	1	0	0	1
4	0	1	0.283	0.931	0	1	0.060	0.940
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
17	0.500	0	1	0	1	0	1	0
18	2.024	0.340	0	2.128	0.830	0.170	0	1
19	3.846	0	7.143	0	1	0	1	0
20	0	1	1.275	0.725	0	1	0.268	0.732
Mean	1.307	0.179	0.693	0.927	0.812	0.188	0.390	0.610

Table 3: LP Estimates of TFP Change

Row	Firm	Period	q_1	q_2	x_1	x_2	A	M	P
1	1	1	1	1	1	1	1	1	1
2	2	1	0.41	0.5	0.22	0.47	1.159	1.217	0.926
3	3	1	0.13	0.11	0.07	0.18	0.96	1.011	0.764
4	4	1	0.3	1	0.21	1.01	0.576	0.687	0.345
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
17	1	5	2	2	1	1	2	2	2
18	2	5	0.41	0.5	0.22	0.47	1.159	1.217	0.926
19	3	5	0.26	0.22	0.14	0.36	0.96	1.011	0.764
20	4	5	0.3	1	0.21	1.01	0.576	0.687	0.345

A = additive; M = multiplicative; P = primal

5. Least Squares (LS) Estimators

LS methods can also be used to compute additive, multiplicative and primal TFP index numbers. This usually involves the following assumptions:

LS1: all relevant quantities, prices and environmental variables are observed and measured without error;

LS2: the functional forms of distance functions are known.

Assumption LS1 is the same as assumption LP1. If LS1 and LS2 are true, then output and input distance functions can be rewritten in the form of regression models with unobserved errors representing technical inefficiency. The associated models are often referred to as deterministic frontier models (DFMs). The idea behind LS estimation of DFMs can be traced back at least as far as Winsten (1957, p. 283). Aigner and Chu (1968) estimated DFMs using linear and quadratic programming methods. Schmidt (1976) showed that the Aigner and Chu (1968) linear and quadratic programming estimators are ML estimators under the assumption that the inefficiency effects are exponential and half-normal random variables respectively. These alternative estimators are rarely used, not least because their asymptotic properties are unknown: for details, see Schmidt (1976, p. 239) and Greene (1980, Sect. 3.2).

5.1. Output-Oriented DFMs

If LS1 is true, then the linear homogeneity property of ODFs can be used to write $D^t_O(x_{it}, q_{it}, z_{it}) = q_{1it} D^t_O(x_{it}, q^*_{it}, z_{it})$ where $q^*_{it} \equiv q_{it}/q_{1it}$ is a vector of normalised outputs (the choice of normalising output is arbitrary). Equivalently,

$$\ln q_{1it} = -\ln D^t_O(x_{it}, q^*_{it}, z_{it}) - u_{it} \tag{25}$$

where $u_{it} \equiv -\ln D^t_O(x_{it}, q_{it}, z_{it}) \geq 0$ is an output-oriented technical inefficiency effect (the value of the ODF is a measure of output-oriented technical efficiency). Assumption LS2 says the functional form of the ODF is known. If the ODF is given by (5), for example, then

$$\ln q_{1it} = \alpha(t) + \delta \ln z_{it} + \sum_{m=1}^{M} \beta_m \ln x_{mit} - \ln \left(\sum_{n=1}^{N} \gamma_n q^*_{nit} \right) - u_{it} \tag{26}$$

where $\alpha(t) = \ln A(t)$, $\alpha(t) \geq \alpha(t-1)$, $\beta = (\beta_1, \ldots, \beta_M)' \geq 0$, $\gamma = (\gamma_1, \ldots, \gamma_N)' \geq 0$ and $\sum_n \gamma_n = 1$. If there is only one output and there is no technical or environmental change, then this function collapses to a DFM that has the same basic structure as the Cobb-Douglas models of Aigner and Chu (1968, p. 831), Richmond (1974, p. 517) and Schmidt (1976, Eq. 2).

5.2. *Input-Oriented DFMs*

If LS1 is true, then the linear homogeneity property of IDFs can be used to write $D_I^t(x_{it}, q_{it}, z_{it}) = x_{1it} D_I^t(x_{it}^*, q_{it}, z_{it})$ where $x_{it}^* \equiv x_{it}/x_{1it}$ is a vector of normalised inputs (again, the choice of normalising input is arbitrary). Equivalently,

$$-\ln x_{1it} = \ln D_I^t(x_{it}^*, q_{it}, z_{it}) - u_{it} \tag{27}$$

where, in a slight abuse of notation, $u_{it} \equiv \ln D_I^t(x_{it}, q_{it}, z_{it}) \geq 0$ is now an input-oriented technical inefficiency effect (the reciprocal of the value of the IDF is a measure of input-oriented technical efficiency). Again, LS2 says the functional form of the IDF is known. If the IDF is given by (6), for example, then

$$-\ln x_{1it} = \xi(t) + \kappa \ln z_{it} + \sum_{m=1}^{M} \lambda_m \ln x_{mit}^* - \frac{1}{\eta} \ln \left(\sum_{n=1}^{N} \gamma_n q_{nit} \right) - u_{it} \tag{28}$$

where $\xi(t) = \ln B(t)$. If there is only one output, then this function collapses to a Cobb-Douglas DFM.

5.3. *Estimation*

Irrespective of the orientation of the DFM, it is common to assume that u_{it} is a random variable with the following properties:

LS3 $E(u_{it}) = \mu \geq 0$ for all i and t;

LS4 $var(u_{it}) \propto \sigma_u^2$ for all i and t;

LS5 $cov(u_{it}, u_{ks}) = 0$ if $i \neq k$ or $t \neq s$; and

LS6 u_{it} is not correlated with the observed explanatory variables.

If LS1–LS6 are true, then LS estimators for the unknown parameters in DFMs are consistent. Different types of restricted and unrestricted LS estimators must generally be used to estimate different types of DFMs. For example, consider the output-oriented DFM given by (26). If LS1 is true, then this model can be rewritten as

$$\ln q_{1it} = \alpha^*(t) + \delta \ln z_{it} + \sum_{m=1}^{M} \beta_m \ln x_{mit} - \ln \left(\sum_{n=1}^{N} \gamma_n q^*_{nit} \right) - e_{it} \tag{29}$$

where $\alpha^*(t) = \alpha(t) - \mu$, $\alpha(t) \geq \alpha(t-1)$, $\beta = (\beta_1, \ldots, \beta_M)' \geq 0$, $\gamma = (\gamma_1, \ldots, \gamma_N)' \geq 0$, $\sum_n \gamma_n = 1$ and $e_{it} = \mu - u_{it}$. This model is nonlinear in the parameters. In this case, the *restricted nonlinear* least squares (RNLS) estimator is a consistent estimator for the slope coefficients. A consistent estimator for μ is $\hat{\mu} = \max\{\hat{e}_{11}, \ldots, \hat{e}_{IT}\}$ where $\hat{e}_{it}$ is the it-th RNLS residual. A consistent estimator for $\alpha(t)$ is $\hat{\alpha}(t) = \hat{\alpha}^*(t) + \hat{\mu}$ where $\hat{\alpha}^*(t)$ is the RNLS estimator for $\alpha^*(t)$. In this paper, these estimators are collectively referred to as *corrected* restricted nonlinear least squares (CRNLS) estimators.

5.4. *TFP Indices*

LS methods can be used to compute additive, multiplicative and primal TFP index numbers. The precise nature of the computations depends on the distance functions. If the ODF and IDF are given by (5) and (6), for example, then additive, multiplicative and primal TFP index numbers can be computed using (12), (13) and (14). Estimates of the unknown parameters in these equations can be obtained by estimating either (26) or (28) [a useful property of DFMs is that the (estimated) parameters of any functional representation of a technology set can, in principle, always be recovered from the (estimated) parameters of any other functional representation of the technology set; thus, in principle, it doesn't matter if we estimate (26) or (28)].

5.5. *Example*

Reconsider the simulated data reported in Table 1. These data have been used to obtain CRNLS estimates of the unknown parameters in (26). The estimated coefficients and their standard errors are reported in Table 4. The estimates reported in this table are correctly signed by construction.[3] Associated additive, multiplicative and primal index

[3]Observe that the estimates of $\alpha(2)$ and $\alpha(3)$ are exactly zero, and that the standard errors for these coefficients are also exactly zero. This implies that the constraints $\alpha(2) \geq \alpha(1)$ and $\alpha(3) \geq \alpha(2)$ are

numbers are reported in Table 5. Again, the index numbers in any given column of this table provide a coherent picture of TFP change. For example, in any given column, the index number in row 17 correctly indicates that the TFP of firm 1 in period 5 was double what it had been in period 1, and the index numbers in rows 3 and 19 correctly indicate that firm 3 was equally productive in both periods.

Table 4: CRNLS Parameter Estimates

Parameter	Estimate	St. Err.	t-value	p-value
$\alpha(1)$	0.430	0.032	13.613	$<0.01^{***}$
$\alpha(2)$	0	0	NaN	NaN
$\alpha(3)$	0	0	NaN	NaN
$\alpha(4)$	0.407	0.035	11.631	$<0.01^{***}$
$\alpha(5)$	0.403	0.059	6.797	$<0.01^{***}$
δ	-0.328	0.312	-1.053	0.312
β_1	0.680	0.230	2.957	0.011^{**}
β_2	0.273	0.324	0.844	0.414
γ_1	0.822	0.189	4.336	$<0.01^{***}$
γ_2	0.178	0.189	0.941	0.364

***, ** and * indicate significance at 1%, 5% and 10% levels

6. Maximum Likelihood (ML) Estimators

ML estimation of TFP change is underpinned by the assumption that distance functions merely exist. If they exist, then they can be rewritten in the form of regression models with unobserved errors representing inefficiency and statistical noise. The associated models are often referred to as stochastic frontier models (SFMs). ML estimation of SFMs can can be traced back to Aigner et al. (1977) and Meeusen and van den Broeck (1977). The estimated parameters of SFMs can be used to compute additive and multiplicative TFP index numbers. They cannot generally be used to compute primal TFP index numbers: except in restrictive special cases (e.g., there are no functional form errors and all inputs and outputs are observed and measured without error), the relation-

binding. LS methods for imposing inequality constraints are unsatisfactory in this respect. Bayesian methods are much more satisfactory. For details, see Geweke (1986).

Table 5: CRNLS Estimates of TFP Change

Row	Firm	Period	q_1	q_2	x_1	x_2	A	M	P
1	1	1	1	1	1	1	1	1	1
2	2	1	0.41	0.5	0.22	0.47	1.641	1.566	1.558
3	3	1	0.13	0.11	0.07	0.18	1.446	1.366	1.378
4	4	1	0.3	1	0.21	1.01	1.261	1.185	1.29
:	:	:	:	:	:	:	:	:	:
17	1	5	2	2	1	1	2	2	2
18	2	5	0.41	0.5	0.22	0.47	1.641	1.566	1.558
19	3	5	0.26	0.22	0.14	0.36	1.446	1.366	1.378
20	4	5	0.3	1	0.21	1.01	1.261	1.185	1.29

A = additive; M = multiplicative; P = primal

ships between the parameters of SFMs and the components of primal TFP indices are unknown.

6.1. Output-Oriented SFMs

If the ODF exists, then the relationship between inputs, outputs and environmental variables can be written in the form of (25). If the functional form of the ODF is unknown, then (25) can be rewritten as

$$\ln q_{1it} = f^t(x_{it}, q^*_{it}, z_{it}) + v_{it} - u_{it} \tag{30}$$

where $f^t(.)$ is an approximating function chosen by the researcher (e.g., a translog or Cobb-Douglas function), $v_{it} \equiv -\ln D^t_O(x_{it}, q^*_{it}, z_{it}) - f^t(x_{it}, q^*_{it}, z_{it})$ is an unobserved variable that accounts for functional form errors and other sources of statistical noise, and all other variables are as defined in (25). The exact nature of the noise component depends on both the SFM and the unknown ODF. To illustrate, suppose the SFM is

$$\ln q_{1it} = \alpha + \delta \ln z_{it} + \sum_{m=1}^{M} \beta_m \ln x_{mit} - \sum_{n=1}^{N} \xi_n \ln q^*_{nit} + v_{it} - u_{it} \tag{31}$$

where $\sum_n \xi_n = 1$. If the ODF is given by (5), for example, then the noise component in this model is

$$v_{it} = [\ln A(t) - \alpha] + \left[\sum_{n=1}^{N} \xi_n \ln q^*_{nit} - \ln \left(\sum_{n=1}^{N} \gamma_n q^*_{nit} \right) \right]. \tag{32}$$

The first term on the right-hand side can be viewed as an omitted (dummy) variable error. The second term is a functional form error.

6.2. Input-Oriented Models

If the IDF exists, then the relationship between inputs, outputs and environmental variables can be written in the form of (27). If the functional form of the IDF is unknown, then (27) can be rewritten as

$$-\ln x_{1it} = f^t(x^*_{it}, q_{it}, z_{it}) + v_{it} - u_{it} \tag{33}$$

where, in another slight abuse of notation, $f^t(.)$ is an approximating function chosen by the researcher, $v_{it} \equiv \ln D^t_I(x^*_{it}, q_{it}, z_{it}) - f^t(x^*_{it}, q_{it}, z_{it})$ is an unobserved variable that accounts for statistical noise, and all other variables are as defined in (27). Again, the exact nature of the noise component depends on both the SFM and the unknown IDF. To illustrate, suppose the SFM is

$$-\ln x_{1it} = \xi(t) + \kappa \ln z_{it} + \sum_{m=1}^{M} \lambda_m \ln x^*_{mit} + \sum_{n=1}^{N} \theta_n \ln q_{nit} + v_{it} - u_{it}. \tag{34}$$

If the IDF is given by (6), for example, then the noise component in this model is

$$v_{it} = -\frac{1}{\eta} \ln \left(\sum_{n=1}^{N} \gamma_n q_{nit} \right) - \sum_{n=1}^{N} \theta_n \ln q_{nit}. \tag{35}$$

This is a functional form error.

6.3. Estimation

Two of the most common assumptions found in the stochastic frontier literature are the following:

ML1 u_{it} is an independent $N^+(0, \sigma^2_u)$ random variable, and

ML2 v_{it} is an independent $N(0, \sigma_v^2)$ random variable.

Here, the term *independent* means, *inter alia*, that v_{it} and u_{it} are not correlated with the other explanatory variables or with each other. Assumption ML1 says that u_{it} is a half-normal random variable obtained by truncating the $N(0, \sigma_u^2)$ distribution from below at zero. This assumption can be traced back to Aigner et al. (1977). Other assumptions concerning the distribution of u_{it} have been made by Meeusen and van den Broeck (1977) (exponential) and Stevenson (1980) (gamma, truncated-normal). If assumptions about the distributions of the error terms in SFMs are true, then ML estimators for the unknown parameters are consistent.

6.4. TFP Indices

ML estimates of the parameters of SFMs can be used to compute additive and multiplicative TFP index numbers, but they cannot generally be used to compute primal TFP index numbers. Computing primal TFP index numbers involves estimating the parameters of output and input distance functions. Except in restrictive special cases (e.g., there are no functional form errors), the parameters of output and input distance functions are not equal to the parameters of the approximating functions in SFMs.

To see how the parameters of approximating functions can be used to compute additive and multiplicative TFP index numbers, suppose the SFM is given by (31). Additive TFP index numbers can be computed using any nonnegative measures of relative value as weights. If the ML estimate of ξ_n in (31) is nonnegative, then it can be used as an output weight. If the ML estimate of β_m is nonnegative, then it can be used as an input weight. On the other hand, multiplicative TFP index numbers can be constructed using any nonnegative measures of value *that sum to one* as weights. If the ML estimates of $\xi_1, \ldots, \xi_N$ in (31) are all nonnegative, then they can be used as output weights. If the ML estimates of $\beta_1, \ldots, \beta_M$ are all nonnegative, then they can be normalised to sum to one. The normalised estimates can then be used as input weights. For a real-world example, see Ding et al. (2016).

6.5. Example

Reconsider the simulated data reported in Table 1. These data have been used to obtain ML estimates of the unknown parameters in (31). The estimated coefficients and their standard errors are reported in Table 6. Associated additive and multiplicative index numbers are reported in Table 7. Again, the index numbers in any given column of this table provide a coherent picture of TFP change. For example, in any given column, the index number in row 17 correctly indicates that the TFP of firm 1 in period 5 was double what it had been in period 1, and the index numbers in rows 4 and 20 correctly indicate that firm 4 was equally productive in both periods.

Table 6: ML Parameter Estimates

Parameter	Estimate	St. Err.	t-value	p-value
α	0.007	0.114	0.062	0.952
δ	-0.143	0.506	-0.282	0.782
β_1	0.425	0.087	4.893	$<0.01^{***}$
β_2	0.610	0.032	18.894	$<0.01^{***}$
ξ_1	0.527	0.018	29.356	$<0.01^{***}$
ξ_2	0.473	0.018	26.338	$<0.01^{***}$

***, ** and * indicate significance at 1%, 5% and 10% levels

Table 7: ML Estimates of TFP Change

Row	Firm	Period	q_1	q_2	x_1	x_2	A	M
1	1	1	1	1	1	1	1	1
2	2	1	0.41	0.5	0.22	0.47	1.232	1.308
3	3	1	0.13	0.11	0.07	0.18	0.894	0.983
4	4	1	0.3	1	0.21	1.01	0.926	1.000
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
17	1	5	2	2	1	1	2	2
18	2	5	0.41	0.5	0.22	0.47	1.232	1.308
19	3	5	0.26	0.22	0.14	0.36	0.894	0.983
20	4	5	0.3	1	0.21	1.01	0.926	1.000

A = additive; M = multiplicative

7. Other Measures of TFP Change

LP methods are often used to estimate binary TFP indices (i.e., indices for making comparisons involving only two observations). LS methods are often used to estimate residual measures of TFP change. Except in restrictive special cases, binary TFP indices and residual measures of TFP change are not proper TFP indices (among other things, binary TFP indices do not generally satisfy the transitivity axiom T7, and residual measures of TFP change do not generally satisfy the proportionality axiom T5).

7.1. Binary TFP Indices

On the output side, a generalised Malmquist (GM) index that compares the outputs of firm i in period t with the outputs of firm k in period s is

$$QI^{GM}(x_{ks}, x_{it}, q_{ks}, q_{it}, z_{it}, z_{ks}, s, t) = \left(\frac{D_O^t(x_{it}, q_{it}, z_{it})}{D_O^t(x_{it}, q_{ks}, z_{it})} \frac{D_O^s(x_{ks}, q_{it}, z_{ks})}{D_O^s(x_{ks}, q_{ks}, z_{ks})} \right)^{1/2}. \qquad (36)$$

If there is only one firm involved in the comparison, then the i and k subscripts can be suppressed. If there is only one firm and there is no environmental change, then the two ratios in this index are equal to the Malmquist output indices defined by Diewert (1992, Eqs. 118, 119) (hence the term "generalised Malmquist"). On the input side, a GM index that compares the inputs of firm i in period t with the inputs of firm k in period s is

$$XI^{GM}(x_{ks}, x_{it}, q_{ks}, q_{it}, z_{it}, z_{ks}, s, t) = \left(\frac{D_I^t(x_{it}, q_{it}, z_{it})}{D_I^t(x_{ks}, q_{it}, z_{it})} \frac{D_I^s(x_{it}, q_{ks}, z_{ks})}{D_I^s(x_{ks}, q_{ks}, z_{ks})} \right)^{1/2}. \qquad (37)$$

If there is only one firm and there is no environmental change, then the two ratios in this index are equal to the Malmquist input indices defined by Diewert (1992, Eqs. 97, 98). Finally, the GM index that compares the TFP of firm i in period t with the TFP of firm k in period s is

$$TFPI^{GM}(x_{ks}, x_{it}, q_{ks}, q_{it}, z_{it}, z_{ks}, s, t) = \frac{QI^{GM}(x_{ks}, x_{it}, q_{ks}, q_{it}, z_{it}, z_{ks}, s, t)}{XI^{GM}(x_{ks}, x_{it}, q_{ks}, q_{it}, z_{it}, z_{ks}, s, t)}. \qquad (38)$$

The linear homogeneity properties of output and input distance functions mean that this index can also be written in terms of normalised shadow output and input prices. If normalised shadow prices are proportional to observed market prices, then, and only then, the GM TFP index is equivalent to the well-known Fisher TFP index.

In the productivity literature, many binary TFP indices are defined using functional representations of *period-specific* production possibilities sets. A period-specific production possibilities set is a set containing all input-output combinations that are physically possible *in a given period*. Mathematically, the set of input-output combinations that are possible in period t is $T^t = \{(x,q) : x \text{ can produce } q \text{ in period } t\}$. If assumptions (b), (d) and (e) from Section 2.1 are true, then this set can be represented by *period-specific* distance functions. In this paper, the period-t ODF and IDF are defined as

$$D_O^t(x_{it}, q_{it}) = \inf\{\rho > 0 : (x_{it}, q_{it}/\rho) \in T^t\} \tag{39}$$

$$\text{and} \qquad D_I^t(x_{it}, q_{it}) = \sup\{\theta > 0 : (x_{it}/\theta, q_{it}) \in T^t\}. \tag{40}$$

These distance functions can be used to define three well-known binary indices.

First, the Hicks-Moorsteen (HM) index that compares the TFP of firm i in period t with the TFP of firm k in period s is

$$TFPI^{HM}(x_{ks}, x_{it}, q_{ks}, q_{it}, s, t) = \left[\frac{D_O^t(x_{it}, q_{it})}{D_O^t(x_{it}, q_{ks})} \frac{D_O^s(x_{ks}, q_{it})}{D_O^s(x_{ks}, q_{ks})} \frac{D_I^t(x_{ks}, q_{it})}{D_I^t(x_{it}, q_{it})} \frac{D_I^s(x_{ks}, q_{ks})}{D_I^s(x_{it}, q_{ks})} \right]^{1/2}. \tag{41}$$

This type of index was vaguely conceptualised by Hicks (1961) and more clearly conceptualised by Moorsteen (1961) (hence the term “Hicks-Moorsteen”). If there is only one firm and $s = t - 1$, then it is equal to the Malmquist productivity index defined by Bjurek (1996, Eq. 9).

Second, the output-oriented Malmquist (OM) index that compares the TFP of firm i in period t with the TFP of firm k in period s is

$$TFPI^{OM}(x_{ks}, x_{it}, q_{ks}, q_{it}, s, t) = \left[\frac{D_O^s(x_{it}, q_{it})\, D_O^t(x_{it}, q_{it})}{D_O^s(x_{ks}, q_{ks})\, D_O^t(x_{ks}, q_{ks})} \right]^{1/2}. \tag{42}$$

If there is only one firm and $s = t - 1$, then this index is equal to the OM productivity index defined by Färe, Grosskopf, Norris and Zhang (1994, Eq. 6) and Grifell-Tatjé and Lovell (1995, Eq. 8).

Finally, the input-oriented Malmquist (IM) index that compares the TFP of firm i in period t with the TFP of firm k in period s is

$$TFPI^{IM}(x_{ks}, x_{it}, q_{ks}, q_{it}, s, t) = \left[\frac{D_I^s(x_{ks}, q_{ks})\, D_I^t(x_{ks}, q_{ks})}{D_I^s(x_{it}, q_{it})\, D_I^t(x_{it}, q_{it})} \right]^{1/2}. \tag{43}$$

If there is only one firm and $s = t - 1$, then this index is equal to to *the reciprocal of* the input-based Malmquist productivity index defined by Färe et al. (1992, Eq. 9). If inputs are strongly disposable, then IDFs are nondecreasing in inputs. In this case, if a firm uses more inputs to produce the same outputs, then the index defined by (43) will take a value less than or equal to 1, indicating that productivity has fallen or remained the same. However, the index defined by Färe et al. (1992, Eq. 9) will take a value greater than or equal to 1.

The components of GM, HM, OM and IM TFP indices can be computed/estimated using either LP or LS estimators. The LP estimators are simple variants of the linear programs described in Section 4. For example, consider the following variant of problem (16):

$$\min_{\alpha_{it},\delta_{it},\beta_{it},\gamma_{it}} \Big\{ \alpha_{it} + \beta_{it}' x_{it} : \gamma_{it} \geq 0,\ \beta_{it} \geq 0,\ \gamma_{it}' q_{it} = 1,$$
$$\alpha_{it} + \beta_{it}' x_{hr} - \gamma_{it}' q_{hr} \geq 0 \text{ for all } h \leq I \text{ and } r \leq t \Big\}. \tag{44}$$

The optimised value of the objective function is an estimate of $1/D_O^t(x_{it}, q_{it})$. As another example, consider the following variant of problem (19):

$$\max_{\xi_{it},\kappa_{it},\phi_{it},\theta_{it}} \Big\{ \xi_{it}' q_{it} - \phi_{it} : \xi_{it} \geq 0,\ \theta_{it} \geq 0,\ \theta_{it}' x_{it} = 1,$$
$$\xi_{it}' q_{hr} - \phi_{it} - \theta_{it}' x_{hr} \leq 0 \text{ for all } h \leq I \text{ and } r \leq s \Big\}. \tag{45}$$

In this case, the optimised value of the objective function is an estimate of $1/D_I^s(x_{it}, q_{it})$.

7.2. Residual Measures

It is common to use econometric methods to estimate the parameters of aggregate production function models, and to then use changes in the residuals as measures of TFP change. This approach to measuring productivity change has its antecedents in the work of Solow (1957). For a simple example, consider the following production function model:

$$\ln Q_{it} = \alpha + \sum_{m=1}^{M} \beta_m \ln x_{mit} + e_{it} \tag{46}$$

where $Q_{it} \equiv Q(q_{it})$ is an aggregate output and e_{it} is an unobserved variable that accounts for statistical noise. This equation has the same structure as the multiple regression model (MRM) discussed in introductory econometics textbooks. As usual, the exact form of the noise component depends on the form of the ODF. If the ODF is given by (5), for example, then the noise component in (46) is

$$e_{it} = [\ln A(t) - \alpha + \delta \ln z_{it} - u_{it}] + \left[\ln Q(q_{it}) - \ln \left(\sum_{n=1}^{N} \gamma_n q_{nit} \right) \right] \tag{47}$$

where $u_{it} \equiv -\ln D_O^t(x_{it}, q_{it}, z_{it})$ is the usual output-oriented technical inefficiency effect. The first term in brackets on the right-hand side of (47) can be viewed as an omitted variable error. The second term is a (possible) measurement error. To estimate the model, it is common to assume that e_{it} is a random variable with the following properties:

MRM1 $E(e_{it}) = 0$ for all i and t;

MRM2 $var(e_{it}) = \sigma_e^2$ for all i and t;

MRM3 $cov(e_{it}, e_{ks}) = 0$ if $i \neq k$ or $t \neq s$; and

MRM4 $cov(e_{it}, \ln x_{mks}) = 0$ for all i, t, m, k and s.

If MRM1–MRM4 are true, then the ordinary least squares (OLS) estimators for the unknown parameters are unbiased and consistent. The associated residual (R) index that compares the TFP of firm i in period t with the TFP of firm k in period s is

$$TFPI^R(x_{ks}, q_{ks}, x_{it}, q_{it}) = \left(\frac{Q_{it}}{Q_{ks}} \right) \prod_{m=1}^{M} \left(\frac{x_{mks}}{x_{mit}} \right)^{b_m} \tag{48}$$

where b_m is the OLS estimator for β_m. The term "residual" derives from the fact that $TFPI^R(x_{ks}, q_{ks}, x_{it}, q_{it}) = \exp(\hat{e}_{it} - \hat{e}_{ks})$ where $\hat{e}_{it}$ is the it-th OLS residual. Importantly, $TFPI^R(x_{ks}, q_{ks}, x_{it}, q_{it})$ is a proper TFP index *if and only if* $b_m \geq 0$ and $\sum_m b_m = 1$. In practice, this adding-up requirement is often ignored: see, for example, Olley and Pakes (1996, pp. 1273,1287) and Harris and Moffat (2015, Eq. 2).

7.3. Example

Reconsider the simulated data reported in Table 1. These data have been used to compute the GM, HM, OM and IM index numbers reported in Table 8. The GM index numbers were computed using the estimated normalised shadow prices reported in Table 2. The HM, OM and IM numbers were computed using variants of the linear programs described in Section 4 [including problems (44) and (45)]. The index numbers in any given column of Table 8 provide an almost totally incoherent picture of TFP change. For example, the index number in row 17 of the GM column indicates that firm 1 was just as productive period 5 as it had been in period 1, even though it used the same inputs to produce twice as much output; the index numbers in rows 2 and 18 of the HM column indicate that firm 2 was more productive in period 5 than it had been in period 1, even though it used exactly the same inputs to produce exactly the same outputs; the index numbers in rows 3 and 19 of the OM column indicate that firm 3 was more productive in period 5 than it had been in period 1, even though it used exactly twice as much input to produce exactly twice as much output; and the index numbers in rows 4 and 20 of the IM column indicate that firm 4 was more productive in period 5 than it had been in period 1, even though it used exactly the same inputs to produce exactly the same outputs. Observe that the index number in row 17 of the IM column is zero; this is due to the fact that there was no feasible solution to problem (45) for firm 1 in period 5.

Finally, OLS estimates of the unknown parameters in (46) are reported in Table 9.[4] Associated residual measures of TFP change are reported in Table 10. Again, these measures are incoherent. For example, the index numbers in rows 3 and 19 indicate that firm 3 was more productive in period 5 than it had been in period 1, even though it used exactly twice as much input to produce exactly twice as much output. This incoherency is due to the fact that the OLS estimates of the slope coefficients only sum to 0.852. The TFP changes reported in Table 9 are nothing more than changes in statistical noise (i.e, functional form errors, omitted variable errors and/or measurement errors).

8. Conclusion

Total factor productivity (TFP) indices are measures of physical output change divided by measures of physical input change. In this paper, the focus has been on TFP

[4]For illustrative purposes, the aggregate output was computed as $Q(q_{it}) = \hat{p}'q_{it}$ where $\hat{p} = (1.307, 0.179)'$. These weights are the average estimated normalised shadow output prices reported in the last line of Table 2. Thus, the aggregate output is an additive output index that has been constructed using average LP estimates of normalised shadow output prices as weights.

Table 8: LP Estimates of TFP Change

Row	Firm	Period	q_1	q_2	x_1	x_2	GM	HM	OM	IM
1	1	1	1	1	1	1	1	1	1	1
2	2	1	0.41	0.5	0.22	0.47	$1^{\dagger}$	$1.408^{\dagger}$	$1^{\dagger}$	$1^{\dagger}$
3	3	1	0.13	0.11	0.07	0.18	$0.853^{\dagger}$	$1.065^{\dagger}$	$1^{\dagger}$	$1^{\dagger}$
4	4	1	0.3	1	0.21	1.01	$3.312^{\dagger}$	$1.365^{\dagger}$	$1^{\dagger}$	$1^{\dagger}$
:	:	:	:	:	:	:	:	:	:	
17	1	5	2	2	1	1	$1^{\dagger}$	2	2	$0^{\dagger}$
18	2	5	0.41	0.5	0.22	0.47	$0.900^{\dagger}$	$1.652^{\dagger}$	$1.406^{\dagger}$	$1.322^{\dagger}$
19	3	5	0.26	0.22	0.14	0.36	$2.194^{\dagger}$	$1.392^{\dagger}$	$1.384^{\dagger}$	$1.302^{\dagger}$
20	4	5	0.3	1	0.21	1.01	$5.453^{\dagger}$	$1.854^{\dagger}$	$1.414^{\dagger}$	$1.329^{\dagger}$

GM = generalised Malmquist; HM = Hicks-Moorsteen; OM = output-oriented Malmquist;
IM = input-oriented Malmquist; † Incoherent.

Table 9: OLS Parameter Estimates

Parameter	Estimate	St. Err.	t-value	p-value
α	0.142	0.053	2.667	0.016 **
β_1	0.697	0.078	8.886	$<$0.01 ***
β_2	0.155	0.115	1.345	0.196

***, ** and * indicate significance at 1%, 5% and 10% levels

Table 10: Residual Measures of TFP Change

Row	Firm	Period	Q	x_1	x_2	R
1	1	1	1	1	1	1
2	2	1	0.421	0.22	0.47	1.359
3	3	1	0.128	0.07	0.18	$1.062^{\dagger}$
4	4	1	0.384	0.21	1.01	1.139
:	:	:	:	:		
17	1	5	2	1	1	2
18	2	5	0.421	0.22	0.47	1.359
19	3	5	0.255	0.14	0.36	$1.177^{\dagger}$
20	4	5	0.384	0.21	1.01	1.139

R = residual; † Incoherent.

indices that are “proper” in the sense that the component output and input indices satisfy the common notions, or axioms, listed in O’Donnell (2016). The class of proper TFP indices includes several additive, multiplicative and primal indices. This paper has shown how additive and multiplicative index numbers can be computed using linear programming (LP), least squares (LS) and maximum likelihood (ML) methods. It has also shown how primal index numbers can be computed using either LP or LS methods. Indices that are not transitive and are therefore not proper include Malmquist and Hicks-Moorsteen TFP indices. This paper has shown how generalised Malmquist and Hicks-Moorsteen index numbers can be computed using LP methods. Measures of productivity change that do not generally satisfy a proportionality axiom and are therefore not proper include a residual index of the type used by Olley and Pakes (1996). This paper has shown how residual index numbers can be computed using LS methods.

Finally, many researchers use output and input quantity indices that are not proper indices. For these researchers, some final remarks about measurement are warranted. The starting point is to recognise that measurement involves assigning numbers to objects. However, the assignment cannot be done in an arbitrary way. In the representational theory of measurement (RTM), a measure is valid if and only if the relationships found in the objects are also found in the numbers. For example, if object A is three times heavier than object B, then a measure of weight should assign a number to object A that is three times bigger than the number assigned to object B; and if basket A contains the same amount of every good as basket B, then a measure of quantity should assign a number to basket A that is the same as the number assigned to basket B. According to Tal (2016), the RTM has been the most influential mathematical theory of measurement to date. It is the theory of measurement that underpins this paper. However, it is evidently not widely accepted in the index number literature: if basket A contains the same amount of every good as basket B, then the measures of quantity that are peddled most widely in the index number literature, the Fisher and Törnqvist indices, will typically assign different numbers to the two baskets. These differences can be viewed as measurement errors. Importantly, they are avoidable.

References

Aczel, J. and Eichhorn, W. (1974), 'A note on additive indices', *Journal of Economic Theory* **8**(4), 525–529.

Aigner, D. and Chu, S. (1968), 'On estimating the industry production function', *The American Economic Review* **58**(4), 826–839.

Aigner, D., Lovell, C. and Schmidt, P. (1977), 'Formulation and estimation of stochastic frontier production function models', *Journal of Econometrics* **6**, 21–37.

Balk, B. (1998), *Industrial Price, Quantity, and Productivity Indices: The Micro-Economic Theory and an Application*, Kluwer Academic Publishers, Boston.

Balk, B. (2008), *Price and Quantity Index Numbers: Models for Measuring Aggregate Change and Difference*, Cambridge University Press, New York.

Banker, R., Charnes, A. and Cooper, W. (1984), 'Some models for estimating technical and scale inefficiencies in data envelopment analysis', *Management Science* **30**(9), 1078–1092.

Beavis, B. and Dobbs, I. (1990), *Optimization and Stability Theory for Economic Analysis*, Cambridge University Press.

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

Charnes, A., Cooper, W. and Rhodes, E. (1978), 'Measuring the efficiency of decision-making units', *European Journal of Operational Research* **2**(6), 429–444. See also "Corrections" op.cit. 3(4):339.

Coelli, T., Rao, D., O'Donnell, C. and Battese, G. (2005), *An Introduction to Efficiency and Productivity Analysis*, 2nd edn, Springer, New York.

Diewert, W. (1992), 'Fisher ideal output, input, and productivity indexes revisited', *Journal of Productivity Analysis* **3**(3), 211–248.

Ding, S., Guariglia, S. and Harris, R. (2016), 'The determinants of productivity in Chinese large and medium-sized industrial firms, 1998–2007', *Journal of Productivity Analysis* **45**(2), 131–155.

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

Färe, R., Grosskopf, S. and Lovell, C. (1994), *Production Frontiers*, Cambridge University Press, Cambridge.

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

Färe, R. and Primont, D. (1995), *Multi-output Production and Duality: Theory and Applications*, Kluwer Academic Publishers, Boston.

Farrell, M. (1957), 'The measurement of productive efficiency', *Journal of the Royal Statistical Society, Series A (General)* **120**(3), 253–290.

Farrell, M. and Fieldhouse, M. (1962), 'Estimating efficient production functions under increasing returns to scale', *Journal of the Royal Statistical Society, Series A (General)* **125**(2), 252–267.

Fried, H., Lovell, C. A. K. and Schmidt, S. (2008), Efficiency and productivity, *in* H. Fried, C. A. K. Lovell and S. Schmidt, eds, 'The Measurement of Productive Efficiency and Productivity Growth', Oxford University Press, New York, pp. 3–91.

Geweke, J. (1986), 'Exact inference in the inequality constrained normal linear regression model', *Journal of Applied Econometrics* **1**(2), 127–141.

Greene, W. (1980), 'Maximum likelihood estimation of econometric frontier functions', *Journal of Econometrics* **13**(1), 27–56.

Grifell-Tatjé, E. and Lovell, C. (1995), 'A note on the Malmquist productivity index', *Economics Letters* **47**, 169–175.

Harris, R. and Moffat, J. (2015), ‘Plant-level determinants of total factor productivity in Great Britain, 1997–2008’, *Journal of Productivity Analysis* **44**(1), 1–20.

Hicks, J. (1961), *Measurement of Capital in Relation to the Measurement of Other Economic Aggregates*, Macmillan, London.

Hill, P. (2008), Lowe indices, Paper presented at the 2008 World Congress on National Accounts and Economic Performance Measures for Nations, Washington, D.C.

IMF (2004), *Producer Price Index Manual: Theory and Practice*, published for the ILO, IMF, OECD, UNECE and World Bank by the International Monetary Fund, Washington, D.C.

Kumbhakar, S. and Lovell, C. (2000), *Stochastic Frontier Analysis*, Cambridge University Press.

Lowe, J. (1823), *The Present State of England in Regard to Agriculture, Trade and Finance*, 2nd edn, Longman, Hurst, Rees, Orme and Brown, London.

Meeusen, W. and van den Broeck, J. (1977), ‘Efficiency estimation from Cobb-Douglas production functions with composed error’, *International Economic Review* **18**(2), 435–444.

Moorsteen, R. (1961), ‘On measuring productive potential and relative efficiency’, *The Quarterly Journal of Economics* **75**(3), 151–167.

O’Donnell, C. (2012), ‘Nonparametric estimates of the components of productivity and profitability change in U.S. agriculture’, *American Journal of Agricultural Economics* **94**(4), 873–890.

O’Donnell, C. (2014), ‘Econometric estimation of distance functions and associated measures of productivity and efficiency change’, *Journal of Productivity Analysis* **41**(2), 187–200.

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

Olley, G. and Pakes, A. (1996), 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica* **64**(6), 1263–1297.

Richmond, J. (1974), 'Estimating the efficiency of production', *International Economic Review* **15**(2), 515–521.

Schmidt, P. (1976), 'On the statistical estimation of parametric frontier production functions', *The Review of Economics and Statistics* **58**(2), 238–239.

Shephard, R. (1953), *Cost and Production Functions*, Princeton University Press, Princeton.

Shephard, R. (1970), *The Theory of Cost and Production Functions*, Princeton University Press, Princeton.

Solow, R. (1957), 'Technical change and the aggregate production function', *Review of Economics and Statistics* **39**(3), 312–320.

Stevenson, R. (1980), 'Likelihood functions for generalized stochastic frontier estimation', *Journal of Econometrics* **13**(1), 57–66.

Tal, E. (2016), Measurement in science, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', winter 2016 edn, Metaphysics Research Lab, Stanford University.

Winsten, C. (1957), 'Discussion on Mr. Farrell's paper', *Journal of the Royal Statistical Society, Series A (General)* **120**(3), 282–84.