

Center



# Discussion Paper

No. 2006–54

## **ALTERNATIVE MEASURES OF TOTAL FACTOR PRODUCTIVITY GROWTH**

By Thijs ten Raa, Victoria Shestalova

May 2006

ISSN 0924-7815

# Alternative measures of Total Factor Productivity growth

Thijs ten Raa and Victoria Shestalova\*

June 2006

## Abstract

The four main approaches to the measurement of total factor productivity (TFP)-growth and its decomposition are (i) Solow's residual analysis, (ii) the Index Number Approach, (iii) Input-Output Analysis (IO), and (iv) Data Envelopment Analysis (DEA). The corresponding measures of TFP growth are based on different assumptions, which we expose and interrelate. The Solow Residual serves as the benchmark for our comparisons. The interrelationships between the alternative measures permit an interpretation of the differences among them. We consolidate the four alternative measures in a common framework.

JEL Classification Numbers: O47, C67

Keywords: TFP, Solow residual, Index numbers, Input-output, DEA

---

\*Thijs ten Raa: Department of Econometrics and Operations Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, the Netherlands. Phone: +31-13-4662365; fax: +31-13-4663280; e-mail: tenRaa@UvT.nl. Victoria Shestalova: Netherlands Bureau for Economic Policy Analysis, P.O. Box 80510, 2508 GM the Hague, the Netherlands. Phone: +31-70-3383446; fax: +31-70-3383350; e-mail: V.Shestalova@cpb.nl.

# 1 Introduction

We review several approaches to the measurement of total factor productivity (TFP)-growth. Our point of departure is the macroeconomic concept of the Solow Residual. We explain its relationship to the alternative measures of TFP growth, including the microeconomic ones. We focus on Index Numbers, Input-Output Analysis (IO) and Data Envelopment Analysis (DEA). The benchmark and the alternative measures are the four main approaches to the measurement of TFP growth encountered in the literature.

The main conceptual difference between these approaches turns out to be the treatment of prices. Traditional productivity indices, including index numbers, rest on the assumption that the observed prices are competitive, so that factors are paid their marginal products. Under this assumption observed value shares are indeed the appropriate weights for the aggregation of the factor productivities into TFP. Frontier approaches, particularly DEA, make no such assumption. Its TFP-growth measure, the so-called Malmquist index, is based on production statistics only and value shares are generated by the shadow prices of the linear program that determines the production possibility frontier. The input-output analytic framework is on either side of the fence. As is well known, IO accounts for intersectoral linkages and yields a TFP measure that is conceptually close to the macro-economic Solow residual. However, IO can accommodate the shadow prices from a general equilibrium model, which moves its TFP-growth measure close to the DEA's.

The main drawback of the assumption of optimizing behavior that provides economic justification to Törnquist and Fisher index numbers is that it rules out inefficiencies. To the contrary, measures of TFP-growth that make no such assumption are capable of ascribing TFP-growth not only to technical change but also to efficiency change. In particular, the Malmquist index (the main frontier analytic TFP-growth measure) is decomposed into these two terms. The technical change component represents a shift of the production frontier and, therefore, resembles the Solow residual measure as defined under the assumption of optimizing behavior. The TFP-growth measure that arises within the IO analytical framework permits a similar decomposition.

The main contribution of this paper is the consolidation of alternative TFP-growth measures in a common theoretical framework. Although the economic literature contains excellent review articles on productivity indices (see, e.g., Diewert, 1992, and Diewert and Nakamura, 2003), to our knowledge this paper is the first to encompass and interrelate all four main measures.

The paper is organized as follows. In section 2 we review the different measures of productivity growth and link them to the Solow residual measure, our benchmark. In section 3 we interrelate the DEA measure and the IO measures. Section 4 concludes.

## 2 Approaches to the measurement of TFP-growth

The economy maps inputs, collected in vector  $x$ , into outputs, collected in vector  $y$  (which may be one-dimensional though). The mapping reflects the production possibilities,  $P$ . More precisely,  $y \in P(x)$ , the set of output vectors producible with input vector  $x$ . Input and output prices are denoted by row vectors  $w$  and  $p$ , respectively. They are exogenous or endogenous, depending on the approach taken to the measurement of productivity. In the endogenous case, the prices are determined by the production possibilities,  $P$ . Since we are interested in TFP growth, all the symbols depend on time, which we indicate by superscription. In particular,  $y^t \in P^t(x^t)$ . If  $y^t$  is one-dimensional, then its rate of growth is  $\hat{y}^t = \frac{dy^t}{dt}/y^t$ . If  $y^t$  is multi-dimensional, then vector  $\hat{y}^t$  is defined component by component. If not only output, but also input is one-dimensional, then *productivity* is defined simply by the output-input ratio,  $y/x$ , and, therefore, productivity growth is  $\widehat{y/x} = \hat{y} - \hat{x}$ . In the multi-dimensional case productivity still captures the difference between output and input growth, but  $\hat{y} - \hat{x}$  is a vector and we must aggregate its components somehow. The aggregation takes place before or after the subtraction, depending on the approach. Furthermore, since the observations are not continuous, the chosen discretization matters. The four main approaches, Solow residual analysis, the Index Number Approach, IO analysis, and DEA, are presented next.

### 2.1 The benchmark: Solow residual

Solow (1957) defined TFP-growth as the rate of growth of real output not accounted for by the growth of the factor inputs and associated it with a shift in technology.

Output  $y$  is a scalar, input  $x$  is the two-dimensional capital-labor vector  $(K, L)$  and the production possibilities are given by:

$$P^t(x) = \{y : y \leq F(K, L, t)\} \quad (1)$$

where  $F(\cdot, \cdot, t)$  is the production function at time  $t$ . In this approach, prices are endogenous. Taking output as the numeraire commodity,  $p = 1$ . Assuming perfect competition, capital and labor are paid their marginal products:  $w_K = \partial F / \partial K$  and  $w_L = \partial F / \partial L$ . TFP-growth,  $\hat{T}$ , is defined as the residual between the output growth rate,  $\hat{y}$ , and the value-shares weighted average of the two input growth rates:

$$\hat{T} = \hat{y} - (\alpha_K \hat{K} + \alpha_L \hat{L}) \quad (2)$$

$$\alpha_K = w_K K / y, \quad \alpha_L = w_L L / y. \quad (3)$$

Under constant returns to scale,  $\alpha_K$  and  $\alpha_L$  sum to unity and we obtain the national income identity,  $w_K K + w_L L = y$ .

Expression (2) has been named the *Solow residual* and referred to as 'the measure of our ignorance', as it is the part of output growth that cannot be explained by the growth of inputs. Under perfect competition it is easy to show that the absence of slack in  $y \leq F(K, L, t)$  yields the following result:

$$\hat{T} = \frac{1}{F} \frac{\partial F}{\partial t}. \quad (4)$$

Hence the Solow residual measures the shift of the production function. See Solow (1957). In the special case of neutral changes (leaving marginal rates of transformation untouched), the aggregate production function is of the form  $A(t)f(K, L)$  – with  $A(t)$  regarded a technical coefficient – and the formula is reduced further to the following short expression:

$$\hat{T} = \hat{A}. \quad (5)$$

In this case technical change is simply the change of the technical coefficient.

Solow's objects of measurement are real output and inputs. Obviously, these macroeconomic concepts require aggregation. Jorgenson and Griliches (1967) argued that the separation of the value of a transaction into price and quantity components is difficult at any practical level of aggregation and that real output and real input are therefore subject to errors of measurement. According to Jorgenson and Griliches, the most important errors arise from incorrect aggregation (using biased estimates for the implicit rental values of capital and labor services) and from

the incorrect accounting for changes in investment and consumption goods prices. Making various adjustments of the US national product accounts (over the 25 years period after World War II), they concluded that “if real product and real factor input were accurately accounted for, the observed rate of growth of total factor productivity was negligible.”<sup>1</sup> However, the paper by Griliches and Jorgenson did not close the discussion on the measurement and explanation of TFP-growth, but rather stimulated it, including research on aggregation methods.

## 2.2 Index numbers

### 2.2.1 Generalization of the Solow residual

The index number approach is a straightforward extension of Solow’s residual analysis to the case of multiple inputs and outputs. If there are more than two inputs, the input growth rate extends to  $\sum \alpha_i \hat{x}_i$  where  $\alpha_i = w_i x_i / y$  and  $w_i = \partial F / \partial x_i$ . If there is also a multiplicity of outputs, the production possibilities can be parametrized by:

$$P^t(x) = \{y : G(y, t) \leq F(x, t)\}. \quad (6)$$

The function of the outputs,  $G(\cdot, t)$ , is assumed to feature constant returns to scale, just like the function of the inputs,  $F(\cdot, t)$ . This formulation is fairly general. It encompasses not only the standard case with independent outputs, but also the case of joint outputs. In formula (6) output and input are separable, but even this can be dispensed with. More precisely, replacement of  $G(y, t) - F(x, t)$  by  $\Phi(x, y, t)$  with  $\Phi(\cdot, \cdot, t)$  featuring constant returns to scale, will not affect our results. Then production is possible if  $\Phi(x, y, t) \leq 0$ , so that the production possibility set is:

$$P^t(x) = \{y : \Phi(x, y, t) \leq 0\}. \quad (7)$$

Under perfect competition,  $\frac{p_j}{w_i} = -\frac{\partial \Phi / \partial y_j}{\partial \Phi / \partial x_i}$ , it is easy to show that the absence of slack and Euler’s theorem yield the following identity between the national product and national income:

$$py = wx, \quad (8)$$

TFP-growth is defined in this more general context as the (Solow) residual between

---

<sup>1</sup>Jorgenson and Griliches (1967) p.250.

the so-called *Divisia indices* of the output and input growth rates:

$$\widehat{T} = \sum \beta_j \widehat{y}_j - \sum \alpha_i \widehat{x}_i, \quad (9)$$

$$\beta_j = \frac{p_j y_j}{p y}, \quad \alpha_i = \frac{w_i x_i}{w x}. \quad (10)$$

Being in terms of quantities, this is the so-called *primal* expression of TFP-growth. Income identity (8) allows us to simplify the definition given by equations (9)-(10) into:

$$\widehat{T} = -\frac{\partial \Phi / \partial t}{\sum (\partial \Phi / \partial y_j) y_j} = \frac{\partial \Phi / \partial t}{\sum (\partial \Phi / \partial x_j) x_j}. \quad (11)$$

Hence the Divisia-based TFP-growth measures the *shift of the production possibility frontier*. The particular case of one output and two inputs considered by Solow is represented by  $\Phi = y - F(K, L, t)$ , under which equation (11) transforms into equation (4).

Jorgenson and Griliches (1967) show by total differentiation of (8) that

$$\widehat{T} = \sum_i \alpha_i \widehat{w}_i - \sum_j \beta_j \widehat{p}_j. \quad (12)$$

This is the *dual* expression of TFP growth, expressing it in prices. A rather trivial yet illuminating rewrite is:

$$\widehat{T} = \sum_i \alpha_i (\widehat{w}_i - \sum_j \beta_j \widehat{p}_j). \quad (13)$$

Here the terms in parentheses are real productivity growth rates of the various factors. It shows that TFP-growth is the value-shares weighted average of all these real productivity growth rates. This representation imputes productivity growth to the various factors of production and thus justifies the name of the residual: total factor productivity growth. For example, in the case of Solow's (1957) aggregate production function with the two inputs capital and labor, TFP-growth is represented as the value-weighted sum of capital productivity growth and labor productivity growth:  $\widehat{T} = \alpha_K \widehat{w}_K + \alpha_L \widehat{w}_L$ .

### 2.2.2 Törnqvist and Fisher indices

The Divisia indices of the output and input growth rates that enter equation (9) must be approximated in discrete time. Alternative discretizations yield different

productivity measures; the Törnqvist and Fisher ideal indices are the most common ones.<sup>2</sup> We illustrate this for the input growth rate, in the simple capital-labor division of Solow:

$$\alpha_K \widehat{K} + \alpha_L \widehat{L} = (w_K K/y) \widehat{K} + (w_L L/y) \widehat{L} = (w_K \frac{dK}{dt} + w_L \frac{dL}{dt}) / (w_K K + w_L L). \quad (14)$$

Here we substituted national income for the national product in the denominator.

Now we may apply the discretization to  $\widehat{K} = \frac{d \ln K}{dt}$  on the left hand side or to  $\frac{dK}{dt}$  on the right hand side of (14) (or to the derivative of any other monotonic transformation of  $K$ ). The difference, however subtle, explains the coexistence of alternative indices.

In the first case,  $\widehat{K} = \frac{d \ln K}{dt}$  on the left hand side of (14) is approximated by  $\ln K^{t+1} - \ln K^t$ . We also have to deal with the weight. In continuous time it is  $\alpha_K$ . Since we consider the capital growth between the periods  $t$  and  $t+1$ , there is some ambiguity. The most common discrete time weight is the arithmetic average of  $\alpha_K^t$  and  $\alpha_K^{t+1}$ . Approximating the other inputs (of  $x$ ) and outputs (of  $y$ ) in the same way, we obtain the *Törnqvist* or *Translog productivity index*<sup>3</sup>,

$$\widehat{T}_T = \sum_j \frac{1}{2} (\beta_j^t + \beta_j^{t+1}) (\ln y_j^{t+1} - \ln y_j^t) - \sum_i \frac{1}{2} (\alpha_i^t + \alpha_i^{t+1}) (\ln x_i^{t+1} - \ln x_i^t). \quad (15)$$

As in equation (10), the weights are  $\beta_j^t = p_j^t y_j^t / p^t y^t$  and  $\alpha_i = w_i^t x_i^t / w^t x^t$ . The nickname Translog index is due to Diewert (1978), who has shown that the approximation is exact for the translog production function.

In the second case,  $\frac{dK}{dt}$  on the right hand side of (14) is approximated by  $\Delta K = K^{t+1} - K^t$  and similarly for  $\frac{dL}{dt}$ . The input growth rate on the right hand side of equation (14) becomes  $(w_K \Delta K + w_L \Delta L) / (w_K K + w_L L)$  and we have to deal with the weights again. For the time being, consider the weights fixed. (Alternative procedures will be detailed shortly.) Since the denominator (national income) is

<sup>2</sup>For example, Christensen en Jorgenson (1970) used the Törnqvist quantity index as a discrete approximation of the Divisia index.

<sup>3</sup>Notice that here and below we use expressions in terms of growth rates, and not in terms of levels. Therefore, our translog index,  $\widehat{T}_T$ , is in fact the growth rate of the translog index as defined in e.g. Diewert (1992). The same holds for the Fisher index,  $\widehat{T}_F$ .



big, a good approximation to this input growth rate is given by:

$$\begin{aligned}\Delta \ln(w_K K + w_L L) &= \ln(w_K K^{t+1} + w_L L^{t+1}) - \ln(w_K K^t + w_L L^t) \\ &= \ln[(w_K K^{t+1} + w_L L^{t+1}) / (w_K K^t + w_L L^t)].\end{aligned}\quad (16)$$

In the bracketed expression on the right hand side we recognize the well-known quantity indices of input growth. For base-year prices we have the Laspeyres quantity index,  $(w_K^t K^{t+1} + w_L^t L^{t+1}) / (w_K^t K^t + w_L^t L^t)$ , and for current prices we have the Paasche index,  $(w_K^{t+1} K^{t+1} + w_L^{t+1} L^{t+1}) / (w_K^{t+1} K^t + w_L^{t+1} L^t)$ . This price ambiguity is customarily resolved by taking the geometric average of the two (as the price is under the ln), which produces the so-called Fisher index. Extending the procedure to the other inputs (in vector  $x$ ) and approximating the growth of output (vector  $y$ ) in the same way, we obtain the general expression for the *Fisher ideal* or *superlative productivity index*:

$$\begin{aligned}\widehat{T}_F &= \frac{1}{2} [\ln(p^t y^{t+1} / p^t y^t) + \ln(p^{t+1} y^{t+1} / p^{t+1} y^t)] \\ &\quad - \frac{1}{2} [\ln(w^t x^{t+1} / w^t x^t) + \ln(w^{t+1} x^{t+1} / w^{t+1} x^t)].\end{aligned}\quad (17)$$

The index is ideal in the sense that it satisfies a number of axioms. (See, e.g., Diewert, 1992.) The nickname 'superlative' originates from Diewert (1976, p.117), who coined this term for an index number that is exact for a flexible functional form. In other words, for a certain functional form (that provides a second order approximation to an arbitrary twice continuously differentiable aggregator function), the index exactly measures the shift in technology. Diewert (1992) has shown that the above index (17) possesses this property.<sup>4</sup>

Although the Törnqvist and Fisher ideal productivity indices are exact for different production functions, most practical time-series applications yield similar numerical values. (See, e.g., Black et al., 2003.) Here we restrict ourselves to highlighting the link between these indices and the Divisia-based productivity index and do not dwell on their detailed properties or economic justification. For these indices Diew-

---

<sup>4</sup>More specifically, Diewert (1992) proofs that under certain parameter restrictions, the Fischer productivity index is exact for a time-dependent revenue function of the following form:  $r^t(p, x) = \sigma_t (p^T A p x^T C x + \alpha^t p \beta^t x p^T B^t x)^{1/2}$ ,  $A = A^T$ ,  $C = C^T$ ,  $t = 0, 1$ , where  $\sigma_t$  is a positive number,  $A$ ,  $C$ , and  $B^t$  are parameter matrices, and  $\alpha^t$  and  $\beta^t$  are parameter vectors.

ert and Nakamura (2003) interrelate the physical and financial concepts of TFP growth, as well as for the Malmquist index, which is the subject of section 2.4.

### 2.3 Input-Output Analysis: Domar aggregation

The IO literature takes the idea of the Solow residual to an economy consisting of sectors which are linked by their outputs in the form of intermediate inputs. Assume that there are  $n$  sectors, each producing a certain commodity and using the other commodities as intermediate inputs. Let  $z_j$  be the gross output of sector  $j$ ,  $p_j$  its price,  $z_{kj}$  the quantity of the intermediate input supplied to sector  $j$  by sector  $k$  (at price  $p_k$ ),  $x_{ij}$  is the quantity of primary input  $i$  engaged in production in sector  $j$  (at price  $w_i$ ). If profits are zero, the financial balances read:

$$p_j z_j = \sum_k p_k z_{kj} + \sum_i w_i x_{ij}, \quad (18)$$

where  $i, j, k = 1, 2, \dots, n$ . The primary inputs are typically capital and labor and the law of one price is assumed to hold. Analogous to equation (9), sectoral TFP-growth,  $\hat{\tau}_j$ , is defined as the *sectoral Solow residual* between the output growth rate,  $\hat{z}_j$ , and the value-shares weighted average of the input growth rates:

$$\hat{\tau}_j = \hat{z}_j - \left( \sum_k p_k z_{kj} \hat{z}_{kj} + \sum_i w_i x_{ij} \hat{x}_{ij} \right) / (p_j z_j). \quad (19)$$

Introducing *technical coefficients*:

$$a_{kj} = z_{kj} / z_j, \quad b_{ij} = x_{ij} / z_j, \quad (20)$$

we may substitute  $\hat{z}_{kj} = \hat{a}_{kj} + \hat{z}_j$  and  $\hat{x}_{ij} = \hat{b}_{ij} + \hat{z}_j$  in equation (19). The  $\hat{z}_j$ -terms cancel in view of the financial balance (18), and we obtain the equivalent expression for total factor productivity growth as a weighted sum of the reductions in technical coefficients:

$$\hat{\tau}_j = - \left( \sum_k \frac{p_k z_{kj}}{p_j z_j} \hat{a}_{kj} + \sum_i \frac{w_i x_{ij}}{p_j z_j} \hat{b}_{ij} \right). \quad (21)$$

In a most explicit way, this confirms Solow's result (5) that the residual measures technical change. The formal embedding of the IO model in the general framework, (11), is, in obvious matrix notation,  $\Phi(x, y, t) = \max\{B(t)[I - A(t)]^{-1}y - x\}$ , where the maximum is taken with respect to (commodity) components. Notice that  $\Phi$  features the Leontief inverse.

Aggregate TFP-growth can be represented as a combination of the sectoral productivity growths. There is a tricky aggregation issue though, which has been analyzed by Domar (1961). The point is that the national product of an economy does not comprise all gross output  $z_j$ , but only the net output  $y_j = z_j - \sum_k z_{jk}$ ; the intermediate inputs do not belong. Indeed, this relationship reconciles (18) with (8).

Suppose, quite realistically, that gross output comprises 40% intermediate input and 60% factor input. Now imagine that intermediate input and factor input are constant, but that gross output grows by 3%. The increase must accrue to net output. However, since net output is only 60% of gross output, the increase is a hefty 5% in terms of net output. In this thought experiment the sectoral Solow residuals will be around 3%, but aggregate TFP-growth is 5%. It turns out that the sectoral rates of TFP-growth must be aggregated to the level of the macro-economy using the value ratios of the sectoral gross outputs to the *net* output of the economy as weights. These so-called *Domar* weights sum to the gross/net output ratio of the economy, 100/60 in our thought experiment.

The Divisia-based TFP-growth was given by the pair of equations (9, 10) or

$$\hat{T} = \left( \sum p_j \frac{dy_j}{dt} - \sum w_i \frac{dx_i}{dt} \right) / wx, \quad (22)$$

where we used the national income identity,  $py = wx$ , in the denominator. Replacing outputs  $y_j$  by the expressions  $z_j - \sum_k z_{jk} = z_j - \sum_k a_{jk} z_k$  and inputs  $x_i$  by  $\sum_j x_{ij} = \sum_j b_{ij} z_j$  (see equation (20) for the technical coefficients), we obtain in matrix notation:

$$\hat{T} = \left[ p \frac{d(z - Az)}{dt} - w \frac{d(Bz)}{dt} \right] / wx. \quad (23)$$

Applying the product rule for differentiation, the  $\frac{dz}{dt}$ -terms cancel in view of the financial balance, (18), and we end up with the aggregate TFP growth of the economy:

$$\hat{T} = \sum \hat{\tau}_j p_j z_j / wx. \quad (24)$$

This is the Domar aggregation rule. TFP-growth is a weighted average of the sectoral Solow residuals. The weights sum to  $pz/wx$ , which is the gross/net output ratio of the economy, in view of the financial balance, (18).

In a way, Domar aggregation corrects the fact that the *direct* measure of TFP growth at the sectoral level expressed by formula (21) neglects that intermediate

inputs are produced by the system. Domar aggregation makes up for this shortcoming at the macro level. Aulin-Ahmavaara (1999) shows how the correction can be done also at the sectoral level, obtaining *effective* rates of sectoral TFP-growth which essentially capture the indirect effects of output price reductions on the productivity in downstream sectors. Another extension is done by Wolff (1985), who distinguishes the value share effect and inter-industry effect, along with the sectoral technical change effect.<sup>5</sup>

Similarly to the previous section, here too a practical issue arises with respect to discretization of the obtained continuous index numbers. See e.g. Dietzenbacher and Los (1998) for further discussion on this.

## 2.4 Data Envelopment Analysis: Malmquist index

So far we have assumed that all input has been used, either by assuming away slack, or by identifying input-output ratios with technical coefficients. The main distinction of DEA, the approach to productivity mostly used in the operations research and management science literature, is that it accounts for (changes in) utilization rates. Its basic tools are the so-called output and input distance functions. The output distance function measures the relative distance between output and the production possibility frontier, where the latter is constructed by enveloping the data. An output distance of 0.8 signifies that output  $y$  is only 80% of what it could be, given the inputs  $x$ . In other words,  $1.25y$  could still be produced with  $x$ . Formally, for any pair of input and output vectors  $(x, y)$  and any time  $t$  the *output distance function* is defined by:

$$D(x, y, t) = \inf\{\theta : y/\theta \in P^t(x)\}. \quad (25)$$

Recall that output vector  $y$  is producible from input vector  $x$  if  $y \in P^t(x)$ . Following Färe and Grosskopf (1996), assume that the sets  $P^t(x)$  are bounded, closed and convex and that  $P^t$  satisfies strong disposability of inputs and constant returns to scale. Similarly, an input distance measures the relative distance between input and the frontier, but under constant returns to scale the two measures are equivalent (the input distance is the inverse of the output distance), so we leave it.<sup>6</sup>

---

<sup>5</sup>Besides, Wolff (1985) treats capital as a produced commodity (similarly to Peterson, 1979) and Aulin-Ahmavaara (1999) treats both labor and capital as produced means of production.

<sup>6</sup>To compute the distance for some observation  $(x, y)$  we have to solve the following problem

In DEA TFP growth is measured by the Malmquist index, which is derived from distance functions:

$$M = \left[ \frac{D(x^{t+1}, y^{t+1}, t)}{D(x^t, y^t, t)} \cdot \frac{D(x^{t+1}, y^{t+1}, t+1)}{D(x^t, y^t, t+1)} \right]^{1/2}. \quad (26)$$

The above representation of the Malmquist index as geometric average of the ratios of distance functions at points  $t$  and  $t+1$  was first proposed by Färe et al. (1989). The index can be decomposed into two economically meaningful sources of TFP growth efficiency change and technical change:

$$M = \frac{D(x^{t+1}, y^{t+1}, t+1)}{D(x^t, y^t, t)} \left[ \frac{D(x^t, y^t, t)}{D(x^t, y^t, t+1)} \frac{D(x^{t+1}, y^{t+1}, t)}{D(x^{t+1}, y^{t+1}, t+1)} \right]^{1/2}$$

If we assume inefficiency away, the first term drops and the Malmquist index features only the technical change component. This component corresponds to technical change as defined by Solow under the assumption of competitive behaviour (no slack and inputs paid marginal products) and measured by the conventional indices such as Divisia, Törnquist en Fisher indices. The following example illustrates this point.

#### 2.4.1 Example

Let us consider the case of one output and neutral technical changes. In this case the technology can be represented by a production function of the following form:

$$y^t = A(t)F(x^t) \quad (27)$$

and the Solow Residual is equivalent to  $\hat{A}$ , which in the discrete case is expressed as

$$\hat{T} = \ln A(t+1) - \ln A(t) = \ln \frac{A(t+1)}{A(t)}. \quad (28)$$

---


$$\begin{aligned} & \inf_{\theta, \lambda \geq 0} \theta \\ \text{s.t. } & -y/\theta + Y^T \lambda \geq 0 \\ & x - X^T \lambda \geq 0 \end{aligned}$$

in which  $X$  and  $Y$  are matrices composed of vector columns of inputs and outputs corresponding to our sample of production units (economies). Alternatively we could use an input distance function, which shows the maximum possible proportional contraction of all inputs still to be able to produce the same amount of output. This would lead to the same measure of efficiency, because input and output distance functions are equivalent under the assumption of constant returns to scale (see Färe and Grosskopf, 1996).

It is easy to see in this special case that the condition of optimizing behavior (no inefficiency) yields the equivalence of the (technical change component of) Malmquist index and (28).<sup>7</sup> In particular, notice that for this production function the output distance function at  $t$  is as follows:

$$\begin{aligned} D^t(x, y) &= \min\{\theta : y/\theta \leq A(t)F(x)\} = \\ &= \min\{\theta : y/A(t)F(x) \leq \theta\} = \frac{y}{A(t)F(x)}. \end{aligned}$$

Substituting this into the formula for the Malmquist index yields:

$$M(x^{t+1}, y^{t+1}, x^t, y^t) = \frac{y^{t+1}}{F(x^{t+1})} \frac{F(x^t)}{y^t}. \quad (29)$$

Assuming inefficiency away, we obtain that output and input in each time  $t$  are related by (27), which yields the result:

$$M(x^{t+1}, y^{t+1}, x^t, y^t) = \frac{A(t+1)}{A(t)}. \quad (30)$$

Hence the technical change component of Malmquist index is equivalent to the Solow residual measure of technical change (28) above.

#### 2.4.2 Results from the literature

The observation demonstrated in the above example holds in a more general case of non-neutral technical changes. In this respect two important results have been established in the literature.

First, Caves et al. (1982) have shown that the Malmquist index (26) becomes a Törnqvist productivity index (15) provided that the distance functions are of translog form with identical second order coefficients, and that the prices are those supporting cost minimization and profit maximization.

Second, Färe and Grosskopf (1992) proved that under the assumption of maximizing behavior the Malmquist index (26) is approximately equal to the Fisher productivity index (17).

These two results provide a link between the conventional Törnqvist and Fisher productivity indices and the Malmquist index, and formulate the conditions for their

---

<sup>7</sup>Färe et al. (1994) provides a similar illustration for the case of a Cobb-Douglas production function.

equivalence. In both cases the assumption that prices support optimizing behavior, such as profit maximization or cost minimization, plays the crucial role. If we impose optimising behavior, all three indices (Törnqvist, Fisher and Malmquist) represent shifts of the production frontier - or ‘technical change’ as defined by Solow - leading to the respective interpretation of the technical change component of the Malmquist index. While the Törnqvist and Fisher indices are defined in terms of observed values, the Malmquist indices use only primary information on inputs and outputs and do not require input prices or output prices in their computation<sup>8</sup>. The explicit price information is replaced by implicit ‘shadow’ price information, derived from the shape of the frontier. (See Coelli and Psarada Rao, 2001.)

The use of a common framework allows us to formalize the relationship between the Malmquist index and the benchmark concept of the Solow residual.

### 2.4.3 Link to Solow residual

If  $(x^t, y^t)$  belongs to the production possibility set at time  $t$ , equation (7) yields  $\Phi(x^t, y^t, t) \leq 0$ . If we expand output  $y^t$  to  $y^t/D(x^t, y^t, t)$  it is still producible from input  $x^t$ , but just:

$$\Phi[x^t, y^t/D(x^t, y^t, t), t] = 0. \quad (31)$$

By constant returns to scale,

$$\Phi[x^t D(x^t, y^t, t), y^t, t] = 0. \quad (32)$$

Now differentiate totally:

$$\sum \frac{\partial \Phi}{\partial x_j} x_j \frac{dD}{dt} + \sum \frac{\partial \Phi}{\partial x_j} \frac{dx_j}{dt} D + \sum \frac{\partial \Phi}{\partial y_j} \frac{dy_j}{dt} + \frac{\partial \Phi}{\partial t} = 0. \quad (33)$$

Divide by either denominator of (11)<sup>9</sup> and rearrange terms:

$$\frac{\sum(\partial \Phi / \partial y_j) dy_j / dt}{\sum(\partial \Phi / \partial y_j) y_j} - \frac{\sum(\partial \Phi / \partial x_j) dx_j / dt}{\sum(\partial \Phi / \partial x_j) x_j} = - \frac{\partial \Phi / \partial t}{\sum(\partial \Phi / \partial y_j) y_j} + \frac{1}{D} \frac{dD}{dt}. \quad (34)$$

<sup>8</sup>Although in theory the Malmquist indices work with physical inputs and outputs, some information on prices can still be necessary in practice. For example, to use capital as input, one have to be able to measure capital. Then observed prices are needed to aggregate over different capital goods.

<sup>9</sup>Here  $\sum(\partial \Phi / \partial x_j) D x_j = - \sum(\partial \Phi / \partial y_j) y_j$ .

In other words,

$$\sum \beta_j \hat{y}_j - \sum \alpha_i \hat{x}_i = -\frac{\partial \Phi / \partial t}{\sum (\partial \Phi / \partial y_j) y_j} + \hat{D}$$

$$\beta_j = \frac{(\partial \Phi / \partial y_j) y_j}{\sum (\partial \Phi / \partial y_j) y_j}, \quad \alpha_i = \frac{(\partial \Phi / \partial x_i) x_i}{\sum (\partial \Phi / \partial x_i) x_i}.$$

The left hand side features the Solow residual between output growth and input growth,  $\hat{T}$ . The first term on the right hand side measures the shift of the production possibility set, as seen in equation (11), and is called *technical change*. The last term measures the shift in the distance to the frontier and is called *efficiency change*.

The link with the DEA literature is established by expressing technical change in terms of the distance function. This is easy. By equation (32), technical change is equal to  $-\frac{\partial D(x^t, y^t, t)}{\partial t} / D$ . Substituting,

$$\hat{T} = \frac{dD(x^t, y^t, t)}{dt} / D - \frac{\partial D(x^t, y^t, t)}{\partial t} / D = \frac{\partial \ln D(x^t, y^t, t)}{\partial x} \frac{dx^t}{dt} + \frac{\partial \ln D(x^t, y^t, t)}{\partial y} \frac{dy^t}{dt} \quad (35)$$

The right hand side measures the *indirect* change in the logarithm of the distance, through the input and output changes. In discrete time this indirect change is approximated by  $\ln D(x^{t+1}, y^{t+1}, \cdot) - \ln D(x^t, y^t, \cdot)$ . It remains to settle at which point of time this difference is evaluated,  $t$  or  $t + 1$ . It is customary to take the average,  $\frac{1}{2}[\ln D(x^{t+1}, y^{t+1}, t) - \ln D(x^t, y^t, t) + \ln D(x^{t+1}, y^{t+1}, t+1) - \ln D(x^t, y^t, t+1)]$   
 $= \ln \left[ \frac{D(x^{t+1}, y^{t+1}, t)}{D(x^t, y^t, t)} \frac{D(x^{t+1}, y^{t+1}, t+1)}{D(x^t, y^t, t+1)} \right]^{1/2}$ .<sup>10</sup> Thus we obtain the expression for *Malmquist productivity index*,

$$\hat{T}_M = \ln M, \quad M = \left[ \frac{D(x^{t+1}, y^{t+1}, t)}{D(x^t, y^t, t)} \frac{D(x^{t+1}, y^{t+1}, t+1)}{D(x^t, y^t, t+1)} \right]^{1/2}, \quad (36)$$

which we have derived from our benchmark definition of TFP growth as Solow residual.

### 3 Synthesis of Input-Output Analysis and DEA

When reviewing the approaches to TFP growth measurement in the previous section, we established their link to the benchmark concept of Solow residual. In this

---

<sup>10</sup>This is analogous to the Fisher index, which is the geometric average of the Laspeyres and Paasche indices.



section we focus on the relation between the measures used by input-output analysis and DEA. As we have discussed above, the overall TFP growth measure that arises in the neoclassical IO framework is conceptually close to the macro-economic Solow residual. Assuming competitive behavior, the prices used in computation are observable prices.

Ten Raa and Mohnen (2001, 2002) augment the neoclassical measure of TFP growth as follows. They apply the traditional formula of the neoclassical growth accounting, but use the shadow prices obtained from the linear program instead of the observable ones. The obtained measure of TFP is based on fundamentals of the economy, similarly to the Malmquist indices. Let us for simplicity consider a closed economy case, with only labor and capital as primary inputs. The underlying linear program is as follows. Given a Leontief technology, Leontief preferences and endowments, the economy expands the final-demand vector  $f \geq 0$  by reallocating inputs among the sectors:

$$\max_{x,c} cef \text{ subject to:} \quad (37)$$

$$\text{material balance constraint: } (I - A)z \geq cf \quad (38)$$

$$\text{factor inputs: } Bz \leq \begin{bmatrix} K \\ L \end{bmatrix} \quad (39)$$

$$\text{non-negativity: } z \geq 0. \quad (40)$$

Here  $e$  is a unit row vector,  $c$  is an expansion factor,  $z$  is gross output as before,  $A$  and  $B$  are matrices of technical coefficients, and scalars  $K$  and  $L$  are the total capital stock and the labor force in the country. Since there are only two primary inputs, capital and labor, matrix  $B$  consists of two row vectors of capital and labor coefficients  $k$  and  $l$ .

The corresponding dual problem is:

$$\min_{p,w} w_K K + w_L L \text{ subject to:} \quad (41)$$

$$-p(I - A) + wB - \sigma = 0, \quad (42)$$

$$pf = ef, \quad (43)$$

$$p \geq 0, \quad w \geq 0, \quad \sigma \geq 0, \quad (44)$$

where  $p$  and  $w = [w_K, w_L]$  are the respective *shadow* prices of outputs and inputs, and  $\sigma$  is slack.

The complementary slackness condition (see, e.g., ten Raa, 2005, for technical detail) gives us  $\sigma z = 0$ ,  $w_K l z = w_K K$ ,  $w_L l z = w_L L$ . Multiplying (42) by  $z$ , we obtain the well-known macroeconomic identity of the national product and national income:

$$p(I - A)z = w_K K + w_L L. \quad (45)$$

The linear program (37) - (40) basically maximizes the expansion factor  $c$ . However, the objective function is normalized so that the value of the final demand at shadow prices is the same as at observable prices, which is reflected also by condition (43) of the dual problem). Similarly to Data Envelopment Analysis, we interpret the inverse of the expansion factor  $1/c$  as efficiency of the economy. The optimal point represents the potential outcome that a multi-sectoral economy could achieve by changing the allocation of production factors across sectors within the economy, given the current technology and preferences.

In accordance with the definition of the Solow residual, we define the TFP growth as the growth of overall final demand minus the growth of aggregate inputs, however, we use the shadow prices to find the value shares:

$$\hat{T} = \frac{\dot{p}f}{pf} - \frac{w_L \dot{L} + w_K \dot{K}}{w_L L + w_K K}. \quad (46)$$

Here a dot denotes the time derivative  $\frac{d}{dt}$ . The above formula can be rearranged as follows:

$$\begin{aligned} \hat{T} &= \frac{\dot{c}pf}{cpf} - \frac{w_L \dot{L} + w_K \dot{K}}{w_L L + w_K K} = \\ &= \frac{-\dot{c}pf + p(\dot{c}f)}{cpf} - \frac{w(Bz)}{cpf} = \\ &\stackrel{(45)}{=} \frac{\dot{c}}{c} - \frac{(p\dot{A} + w\dot{B})z}{cpf}. \end{aligned} \quad (47)$$

The two terms in equation (47) reflect *efficiency change* and *technical change*, the sources of TFP growth also acknowledged in DEA. Efficiency change describes the change in the distance between the observed situation and the potential outcome

that the economy can achieve. Technical change represents the reduction in technical coefficients, which is similar to equation (21) but evaluated at *shadow* prices and the *optimal* gross output levels. Prices that enter this term as weights show the relative importance of technological changes in different sectors. A dual expression can be derived from the national accounting identity:

$$\hat{T} = -\frac{\dot{c}}{c} + \frac{w_K \dot{K} + w_L \dot{L}}{w_L L + w_K K} - \frac{\dot{p}f}{pf}. \quad (48)$$

Similarly to DEA, the new measure of TFP growth that arises in the IO framework encompasses both technical change and efficiency change. However, in contrast to DEA, where the potential for improvement is determined by cross-sectional or intertemporal best practice, in this model the available production technology is assumed to be represented by the observed technical coefficients. Inefficiency stems from the suboptimal allocation of production within the system, or from wasting the resources (not employing the endowed primary inputs in production).<sup>11</sup>

In the case of an open economy, international trade represents another source of TFP growth. An extension of the above model to the case of an open economy allows us incorporate the effect of change in the terms-of-trade. This effect has been considered by ten Raa and Mohnen (2001, 2002) for the case of a small open economy, and by Shestalova (2001) for the case three large open economies. See also Diewert and Morrison (1986) on the effect of international trade on productivity.

## 4 Conclusion

The paper offers a common framework which links the four main approaches to the measurement of TFP growth rates, namely, Solow residual, Index Numbers, Input-Output approach and DEA. Starting with the original approach to productivity growth measurement by Solow (1957), we have demonstrated that Solow residual framework can be generalized to incorporate multiple inputs and outputs as well as the case of intermediate inputs. In addition, this framework can be extended to

---

<sup>11</sup>Strictly speaking, DEA can incorporate other types of inefficiencies as well (for example, non-radial DEA models can account for the presence of a slack). However, we will not discuss those in this particular application, since the standard Malmquist indices based on DEA with constant returns to scale that are typically used for the TFP measurement operate with technical inefficiency.

allow for production and allocative inefficiency, hence, it can also accommodate the DEA-based TFP growth measure.

For all the main indices considered, we review the main results established in the literature that define the conditions under which these indices yield equivalent (or close) TFP growth measures. The condition of optimizing behavior appears to be crucial in this respect. This condition, which lends theoretical support to the conventional Divisia, Törnqvist or Fisher indices, while not required in the case of Malmquist indices, explains the main conceptual difference between the conventional index numbers and the Malmquist indices. This allows the Malmquist indices to incorporate the effect of efficiency change which is neglected by the other indices. Augmenting the standard production function with the inefficiency term defined by a DEA linear program, we derive the Malmquist index from the standard definition of Solow residual as the difference between output growth and input growth.

Input-Output analysis provides a measure of technical changes conceptually close to the conventional Solow Residual, however direct TFP growth rates computed at the sectoral level neglect the effect of technical change on production of intermediate goods. Domar aggregation reconciles the standard macro-economic Solow residual measure with the TFP growth measure that arises within the neoclassical IO growth accounting.

We have shown that the IO-based TFP growth indices can be augmented to factor in efficiency change (and the terms-of-trade effect). This can be done if the observable prices are replaced by shadow prices obtained from the optimization problem. Although, similarly to DEA, the efficiency is interpreted as the potential for boosting the production to reach the production possibility frontier, there is an important difference in the meaning of the frontier in the two models. In DEA the potential is determined by the observable best practice (possibly achieved by other market participants), while in the augmented input-output model it comes from improving allocations of production factors within a multi-sectoral economy.

## References

- [1] Aulin-Ahmavaara, P. (1999) Effective rates of sectoral productivity change, *Economic Systems Research*, 11(4), 349-364.
- [2] Black, M., M. Guy and N. McLellan (2003) Productivity in New Zealand 1988 to 2002, WP 03/06, New Zealand Treasury.
- [3] Caves, D.W., Christensen, L.R., Diewert, W.E. (1982) The economic theory of index numbers and the measurement of input, output and productivity, *Econometrica*, 50(6), 1393-1414.
- [4] Christensen, L.R., Jorgenson, R.W. (1970) U.S. real product and real factor input, 1929-1967, *Review of Income and Wealth*, 16, 19-50.
- [5] Coelli, T., Psarada Rao, D.S. (2001) Implicit value shares in Malmquist TFP index numbers, *CEPA Working paper* No. 4/2001, University of New England, Armidale, Australia.
- [6] Dietzenbacher, E. and Los, B. (1998) Structural decomposition techniques: sense and sensitivity, *Economic System Research*, 10(4), 307-323.
- [7] Diewert, W.E. (1976) Exact and superlative index numbers, *Journal of Econometrics*, 4, 115-146.
- [8] Diewert, W.E. (1992) Fisher ideal output, input and productivity indexes revisited, *Journal of Productivity Analysis*, 3, 211-248.
- [9] Diewert, W.E., Morrison, C. (1986) Adjusting output and productivity indexes for changes in terms of trade, *Economic Journal*, 96, 659-679.
- [10] Diewert, W.E., Nakamura, A. (2003) Index Number Concepts, Measures and Decompositions of Productivity Growth, *Journal of Productivity Analysis*, 19, 127-159.
- [11] Domar, E.D. (1961), On the Measurement of Technological Change, *Economic Journal*, LXXI, 709-729.
- [12] Färe, R., Grosskopf, S.(1992) Malmquist indexes and Fisher ideal indexes, *Economic Journal*, 102(410), 158-160.

- [13] Färe, R., Grosskopf, S.(1996) *Intertemporal Production Frontiers: with Dynamic DEA*, Kluwer Academic Publishers, Boston.
- [14] Färe, R., Grosskopf, S., Lindgren, B., Roos, P. (1989) Productivity development in Swedish hospitals: a Malmquist output index approach, Discussion paper No. 89-3, Southern Illinois University, Illinois.
- [15] Färe, R., Grosskopf, S., Norris, M., Zhang, Z. (1994) Productivity growth, technical progress, and efficiency change in industrialized countries, *American Economic Review*, 84(1), 66-83.
- [16] Jorgenson, D., Griliches, Z. (1967) The explanation of productivity change, *Review of Economic Studies*, 34(3), 249-283.
- [17] Peterson, W. (1979) Total factor productivity in the UK: a disaggregated analysis, in: Patterson, K.D., Schott, K.(eds), *The Measurement of Capital: Theory and Practice*, MacMillan Press, London, UK, 212-225.
- [18] Shestalova, V. (2001) General equilibrium analysis of international TFP growth rates, *Economic Systems Research*, 13(4), 391-402.
- [19] Solow, R.M. (1957) Technical change and the aggregate production function, *The Review of Economics and Statistics*, 39(3), 312-320.
- [20] ten Raa, T. (2005) *The Economics of Input-Output Analysis*, Cambridge University Press, Cambridge, UK.
- [21] ten Raa, T., Mohnen, P. (2002) Neoclassical growth accounting and frontier analysis: a synthesis, *Journal of Productivity Analysis*, 18 (2), 111-128.
- [22] ten Raa, T., Mohnen, P. (2001) The location of comparative advantages on the basis of fundamentals only, *Economic Systems Research*, 13 (1), 93-108.
- [23] Wolff, E.N. (1994) Productivity measurement within an input-output framework, *Regional Science and Urban Economics*, 24, 75-92.