A PARAMETRIC DECOMPOSITION OF A GENERALIZED MALMQUIST-TYPE PRODUCTIVITY INDEX

Luis OREA
University of Oviedo

May 2000

ABSTRACT

This paper provides a parametric method of decomposing a generalized Malmquist-type productivity index into terms related to technical change, technical efficiency changes and returns to scale. The decomposition is based on parametric estimation of translog output-oriented distance functions and draws on the so-called exact index number approach to the derivation of productivity change measures. An empirical application using panel data from Spanish savings banks is included.

Keywords: Malmquist index, distance functions, savings banks.

Luis Orea
Universidad de Oviedo
Departamento de Economía
Avenida del Cristo, s/n
33071 Oviedo (Spain)
E-mail: lorea@econo.uniovi.es

1. Introduction

Overall productivity measures have enjoyed a great deal of interest among researchers analyzing the performance of firms. In the latest years, applications using Malmquist indexes have begun to be very common in productivity analysis. This index has the advantage over traditional Fisher and Törnqvist productivity indexes of decomposing into technical change and technical efficiency change using only quantity information.

A controversial issue here is how to enhance the Malmquist productivity index so that scale efficiency is taken into account. Grifell and Lovell (1995) show that the Malmquist productivity index introduced by Caves, Christensen and Diewert (1982) does not provide an accurate measure of productivity change because it ignores the potential contribution of scale economies to productivity change. In order to avoid this problem, Balk (1999) proposes several measures of productivity change, encompassing not only technical change and technical efficiency change but also scale efficiency change for a multiple-input and multiple-output firm. Grifell and Lovell (1999) address this issue in a different way. They suggest using a generalized Malmquist productivity index, which can be expressed as the product of a conventional Malmquist productivity index and a Malmquist scale index. They demonstrate in the single-input and single-output context that a generalized Malmquist productivity index accurately measures and decomposes productivity change into technical change, technical efficiency change and scale change components. The approach proposed in the present paper can address this issue in a quite simple way.

The Malmquist productivity index can be calculated using non-parametric techniques (DEA) as in Färe, Grosskopf, Lindgren and Roos (1992), or parametric frontier approaches as in Fuentes, Grifell and Perelman (1998). Overall, they show that several distance functions must be previously calculated -using parametric or non-parametric techniques- in order to compute and decompose a Malmquist productivity index.[1]

This paper provides a parametric method of decomposing a generalized Malmquist-type productivity index into terms related to technical change, technical efficiency changes and returns to scale. In contrast to previous works, the decomposition draws on the so-called exact index number approach to the

[1] If some restrictive conditions are satisfied, the Malmquist productivity index can be calculated directly from price and quantity data. Färe and Grosskopf (1992) show that the Fisher index is equal to a Malmquist index under constant returns to scale and competitive cost minimization performance. Caves, Christensen and Diewert (1982) show under similar restrictive conditions that for translog technologies the Törnqvist index is equivalent to the geometric mean of two Malmquist productivity indexes. With non-constant returns to scale, Grifell and Lovell (1999) prove that the geometric mean of two generalized Malmquist productivity indexes is equal to the Törnqvist productivity index.

derivation of productivity change measures. This approach allows us measuring the returns to scale effect on productivity growth using only the estimated parameters of a translog output-oriented distance function with variable returns to scale.

As a by-product, this paper offers a method of decomposing the traditional Törnqvist productivity index without having to approximate concepts that are specified in the continuous time framework. Building on earlier work by Denny, Fuss and Waverman (1981), Bauer (1990) decomposes the standard Divisia productivity index, which is the continuous counterpart of the Törnqvist index. However, this model relies on time derivatives of production or cost functions. Since there are many ways to approximate continuous time derivatives by discrete differences, the Divisia-based models do not lead to an exact formula for the decomposition of the Törnqvist index. The exact index number approach has the advantage over the continuous Bauer's model of leading a formula that is directly suitable for discrete data.

The paper is organized as follows. In Section 2 a generalized Malmquist-type productivity index is decomposed using output-oriented distance functions. Section 3 includes an application to a panel data from Spanish savings banks during the period 1985-98. Section 4 contains a summary and some conclusions.

2. Decomposition of a generalized Malmquist-type productivity index

The Malmquist index is usually defined in terms of distance functions. The principal advantage of distance functions is that they allow the possibility of specifying multiple-input, multiple-output technologies only using data on quantities. On the other hand, it is closely related with efficiency measurement.

Suppose that the firm's technology in period t can be represented by the technology set: [2]

$$T^t = \{(x^t, y^t) : x^t \in \Re^m_+, y^t \in \Re^n_+, x^t \text{ can produce } y^t\} \tag{1}$$

where $y^t=(y_1^t...y_m^t)$ is the output vector and $x^t=(x_1^t...x_n^t)$ is the input vector at time t. The technology set T^t is the list of the technological feasible combinations of inputs and outputs in period t. For each input vector x^t, let $P^t(x^t)$ be the set of feasible outputs vectors y^t that are obtainable from the input vector x^t. Formally,

$$P^t(x^t) = \{ y^t : (x^t, y^t) \in T^t \} \tag{2}$$

Thus, the ouput distance function can then be defined in terms of the output set as:

---

[2] I assume that technology satisfies the axioms listed in Färe and Primont (1995).

$$D_o^t(x^t, y^t) = \min_{\Omega}\{\Omega > 0 : (\frac{y^t}{\Omega}) \in P^t(x^t)\} \qquad (3)$$

The output distance function is defined as the maximum feasible expansion of the output vector with the input vector held fixed. Given a input vector x^t, the value of the output distance function $D^t_o(x^t,y^t)$ puts $y^t/D^t_o(x^t,y^t)$ on the outer boundary of $P^t(x^t)$ and on the ray through y^t. The preceding discussion suggests that the value of the output distance function is less than or equal to one if the output vector y^t is an element of the feasible production set $P^t(x^t)$. The value of the output distance function equals 1 if there is technical efficiency. Hence the output distance function can be viewed as the reciprocal of Farrell's output-oriented measure of technical efficiency. It also follows from the definition of the output distance function that it is linearly homogeneous in outputs.

Assume now that the firm's output distance function follows a translog form:

$$\begin{aligned} \ln D_o(x^t, y^t, t) = & \qquad (4) \\ & \alpha_0 + \sum_{k=1}^{n} \alpha_k \ln x_k^t + \sum_{j=1}^{m} \beta_j \ln y_j^t \\ & + \frac{1}{2}\sum_{k=1}^{n}\sum_{h=1}^{n} \alpha_{kh} \ln x_k^t \ln x_h^t + \frac{1}{2}\sum_{j=1}^{m}\sum_{h=1}^{m} \beta_{jh} \ln y_j^t \ln y_h^t \\ & + \sum_{k=1}^{n}\sum_{j=1}^{m} \gamma_{kj} \ln x_k^t \ln y_j^t \\ & + \psi_0 t + \frac{1}{2}\psi_{00} t^2 + \sum_{k=1}^{n} \xi_{kt} t \ln x_k^t + \sum_{j=1}^{m} \tau_{jt} t \ln y_j^t \end{aligned}$$

Next, observe that the translog distance function can be regarded as a quadratic function in the variables $\ln x^t$, $\ln y^t$ and t. Hence it is possible to apply the so-called Diewert (1976) quadratic identity. Using this identity, the difference between the distance function (4) evaluated at period t and t+1 can be written as

$$\begin{aligned} \ln D_o(t+1) - \ln D_o(t) = & \qquad (5) \\ & \frac{1}{2}\sum_{j=1}^{m}\left[\frac{\partial \ln D_o(t+1)}{\partial \ln y_j} + \frac{\partial \ln D_o(t)}{\partial \ln y_j}\right] \cdot \left(\ln y_j^{t+1} - \ln y_j^t\right) \\ & + \frac{1}{2}\sum_{k=1}^{n}\left[\frac{\partial \ln D_o(t+1)}{\partial \ln x_k} + \frac{\partial \ln D_o(t)}{\partial \ln x_k}\right] \cdot \left(\ln x_k^{t+1} - \ln x_k^t\right) \\ & + \frac{1}{2}\left[\frac{\partial \ln D_o(t+1)}{\partial t} + \frac{\partial \ln D_0(t)}{\partial t}\right] \end{aligned}$$

where $\ln D_O(t)$ represents the distance function evaluated at appropriate values for period t. This equation can also be expressed as

$$\Delta^M \ln Y_{t,t+1} - \Delta^M \ln X_{t,t+1} = \left[\ln D_o(t+1) - \ln D_o(t)\right] - \frac{1}{2}\left[\frac{\partial \ln D_o(t+1)}{\partial t} + \frac{\partial \ln D_0(t)}{\partial t}\right] \quad (6)$$

where the superscript "M" stands, as it is explained below, for Malmquist-type index, and

$$\Delta^M \ln Y_{t,t+1} = \frac{1}{2}\sum_{j=1}^{m}\left[\frac{\partial \ln D_o(t+1)}{\partial \ln y_j} + \frac{\partial \ln D_o(t)}{\partial \ln y_j}\right]\cdot\left(\ln y_j^{t+1} - \ln y_j^t\right)$$

$$\Delta^M \ln X_{t,t+1} = -\frac{1}{2}\sum_{k=1}^{n}\left[\frac{\partial \ln D_o(t+1)}{\partial \ln x_k} + \frac{\partial \ln D_o(t)}{\partial \ln x_k}\right]\cdot\left(\ln x_k^{t+1} - \ln x_k^t\right)$$

The left-hand side of equation (6) can be viewed as an index of total factor productivity, defined in a broad sense as the difference between the weighted average rate of growth of outputs minus the weighted average rate of growth of inputs. [3] The first term on the right-hand side of (6) measures changes in the value of the output distance function from one period to the next. Since the output distance function is the reciprocal of Farrell's output-oriented measure of technical efficiency, this term measures changes in technical efficiency. The second term captures the shift in technology (technical change) between two periods evaluated at two different observed output and input vectors. In this case, the negative sign transforms technical progress (regress) into a positive (negative) value.

In short, we can use equation (6) to measure the contribution to productivity growth of changes in technical efficiency and technical change given information on the value and the proportional shifts in the distance function over time. This information can be obtained by estimating a translog distance function that satisfies any type of returns to scale.

It is noteworthy that the decomposition presented in equation (6) is similar in form to that of a traditional output-based Malmquist productivity index defined as (Caves, Christensen and Diewert, 1982):

$$M_o = \left[\frac{D_o^t(x^{t+1}, y^{t+1})}{D_o^t(x^t, y^t)} \cdot \frac{D_o^{t+1}(x^{t+1}, y^{t+1})}{D_o^{t+1}(x^t, y^t)}\right]^{1/2} \quad (7)$$

---

3 The weights are respectively output distance elasticities and (negative) input distance elasticities. Since the output distance function is decreasing in inputs, the negative sign in the input weights transforms inputs increases (decreases) into a positive (negative) value.

Note that the right-hand side of this index can be rewritten as

$$M_o = \frac{D_o^{t+1}(x^{t+1}, y^{t+1})}{D_o^t(x^t, y^t)} \cdot \left[\frac{D_o^t(x^t, y^t)}{D_o^{t+1}(x^t, y^t)} \cdot \frac{D_o^t(x^{t+1}, y^{t+1})}{D_o^{t+1}(x^{t+1}, y^{t+1})} \right]^{1/2} \tag{8}$$

where the ratio outside the bracket measures changes in technical efficiency, and the geometric mean of the two ratios inside the bracket captures shifts in the production frontier along two rays.

In the same way that equation (6), this last expression decomposes a Malmquist productivity index into the product of a technical efficiency change term and a technical change term.[4] In this sense, the productivity index in the left-hand side of equation (6) can also be viewed as the *parametric* counterpart of a Malmquist-type productivity index when the output distance function is of translog form. Equation (6) also shows how one can obtain from the estimated shape of the frontier surface the weights associated with the Malmquist productivity index, which are implicit using non-parametric techniques.

Further decompositions of productivity growth are also possible. The decomposition above can be first extended to allow for the effect of nonconstant returns to scale. Following Färe and Primont (1995), returns to scale can be computed from the output distance function as follows

$$RTS(t) = \left[-\sum_{k=1}^{n} \partial \ln D_o(t) / \partial \ln x_k \right] - 1 \tag{9}$$

The expression in brackets is the proportional increase in all outputs caused by an increase in all inputs in the same proportion. Therefore, increasing (decreasing) returns to scale are indicated by a value of RTS greater (less) than zero. Using (9) as a measure of returns to scale, equation (5) can be rearranged as

$$\begin{aligned} \Delta^G \ln Y_{t,t+1} - \Delta^G \ln X_{t,t+1} = {} & \left[\ln D_o(t+1) - \ln D_o(t) \right] \\ & - \frac{1}{2}\left[\frac{\partial \ln D_o(t+1)}{\partial t} + \frac{\partial \ln D_0(t)}{\partial t} \right] \\ & + \frac{1}{2}\sum_{k=1}^{n}\left[RTS(t+1) \cdot e_k(t+1) + RTS(t) \cdot e_k(t) \right] \cdot (\ln x_k^{t+1} - \ln x_k^t) \end{aligned} \tag{10}$$

[4] In equation (7), no restrictions are imposed on the returns to scale nature of the reference technology. If the reference technology is defined as having constant returns to scale, the traditional Malmquist productivity index also includes a scale effect (see Färe, Grosskopf, Norris and Zhang, 1994).

where the superscript "G" stands for Generalized Malmquist-type index, and

$$\Delta^G \ln Y_{t,t+1} = \frac{1}{2}\sum_{j=1}^{m}\left[\varepsilon_j(t+1)+\varepsilon_j(t)\right]\cdot\left(\ln y_j^{t+1} - \ln y_j^t\right)$$

$$\Delta^G \ln X_{t,t+1} = \frac{1}{2}\sum_{k=1}^{n}\left[e_k(t+1)+e_k(t)\right]\cdot\left(\ln x_k^{t+1} - \ln x_k^t\right)$$

$$\varepsilon_j(t) = \frac{\partial \ln D_o(t)/\partial \ln y_j}{\sum_{j=1}^{m}\partial \ln D_o(t)/\partial \ln y_j}$$

$$e_k(t) = \frac{\partial \ln D_o(t)/\partial \ln x_k}{\sum_{k=1}^{n}\partial \ln D_o(t)/\partial \ln x_k}$$

The left-hand side of equation (10) is the growth in outputs not accounted by the growth in inputs, where the weights are respectively distance elasticity shares respect to outputs and inputs.[5] The right-hand side of (10) decomposes productivity growth into technical efficiency change and technical change -just like in equation (6)- and the effect of nonconstant returns to scale when inputs expand over time (i.e. movements along the distance function). This last term depends on the degree of local returns to scale and on changes in input quantities. In particular, the scale term vanishes under the assumption of constant returns to scale (or constant input quantities) and, hence, returns to scale have not any effect on productivity growth.

Adding the three terms referred above, the equation (10) provides an accurately measure of productivity change when a translog output distance function with variable returns to scale is estimated using econometric techniques. The attribution of productivity growth to technical efficiency change, technical change, and scale economies in the multiple output case is made without recourse to price information required by Bauer (1990).

It is worth noting that the decomposition presented in equation (10) is quite similar to that of the generalized Malmquist productivity indexes introduced recently by Grifell and Lovell (1999). An output-oriented period t generalized Malmquist productivity index can be expressed and decomposed as follows: [6]

[5] Note that $\Delta^G \ln Y_{t,t+1} = \Delta^M \ln Y_{t,t+1}$ due to the output distance function is homogeneous of degree +1 in the output quantities.

[6] See equation (3.4) in Grifell and Lovell (1999).

$$G_o^t = \frac{D_o^{t+1}(x^{t+1},y^{t+1})}{D_o^t(x^t,y^t)} \cdot \frac{D_o^t(x^{t+1},y^{t+1})}{D_o^{t+1}(x^{t+1},y^{t+1})} \cdot \left[\frac{D_{oc}^t(x^{t+1},y^t)}{D_o^t(x^{t+1},y^t)} \cdot \frac{D_o^t(x^t,y^t)}{D_{oc}^t(x^t,y^t)}\right]^{1/2} \qquad (11)$$

where the subscript "c" in D_o indicates output distance functions defined relative to a reference technology characterized by constant returns to scale. In accordance with this equation, the generalized Malmquist productivity index is decomposed as the product of a technical efficiency change term, a technical change term calculated using period t+1 data, and a scale effect that measures the contribution of scale economies to productivity change.[7]

In summary, the generalized Malmquist productivity index introduced by Grifell and Lovell (1999) includes the effect of scale economies into the measurement of productivity growth, just like the index in the left-hand side of equation (10). In this sense, the right-hand side of equation (10) can be viewed as a parametric decomposition of a generalized Malmquist-type productivity index. This decomposition is based on the estimation of a translog output-oriented distance function with variable returns to scale. Hence, equation (10) provides an exact decomposition for a translog approximation to the structure of production technology.

However, the scale effect terms in both decompositions differ. While the scale effect in Grifell and Lovell (1999) is based on distance functions defined relative to both variable and constant returns to scale reference technologies, the scale effect here is defined relative to distance function with variable returns to scale. So it does not impose any restriction on the returns to scale nature of the reference technology.

---

[7] In this formulation, the technical change term is evaluated using only period t+1 data, and the scale effect is calculated relative to the same period t technology. Alternatively, one could define a generalized Malmquist productivity index using technology in period t+1 as the reference technology. In order to avoid choosing an arbitrary benchmark, one can employ the geometric mean of adjacent-period generalized Malmquist productivity indexes. In this case, the decomposition can be written as:

$$G_O = M_O \cdot \left[\frac{D_{OC}^t(x^{t+1},y^t)/D_O^t(x^{t+1},y^t)}{D_{OC}^t(x^t,y^t)/D_O^t(x^t,y^t)}\right]^{1/2} \cdot \left[\frac{D_{OC}^{t+1}(x^{t+1},y^{t+1})/D_O^{t+1}(x^{t+1},y^{t+1})}{D_{OC}^{t+1}(x^t,y^{t+1})/D_O^{t+1}(x^t,y^{t+1})}\right]^{1/2}$$

where the first term coincide with the traditional Malmquist productivity index defined in equation (8). The last two terms measure the contribution of scale economies to productivity change using the period t and t+1 technology, respectively.

Finally, equation (10) can be used to derive a translog-consistent decomposition of the traditional Törnqvist productivity index. This index can be expressed as:

$$\Delta^{T} \ln TFP_{t,t+1} = \sum_{j=1}^{m} \frac{1}{2}(R_j^t + R_j^{t+1})\left[\ln y_j^{t+1} - \ln y_j^t\right] - \sum_{k=1}^{n} \frac{1}{2}(S_k^t + S_k^{t+1})\left[\ln x_k^{t+1} - \ln x_k^t\right] \qquad (12)$$

where the superscript "T" stands for Törnqvist index, R_j^t is the revenue share of output y_j in total revenue, and S_k^t the cost share of input x_k in total cost. Using this definition, equation (10) can be rearranged as follows

$$\begin{aligned}
\Delta^{T} \ln TFP_{t,t+1} = & \qquad (13)\\
& \left[\ln D_o(t+1) - \ln D_o(t)\right] \\
& - \frac{1}{2}\left[\frac{\partial \ln D_o(t+1)}{\partial t} + \frac{\partial \ln D_0(t)}{\partial t}\right] \\
& + \frac{1}{2}\sum_{k=1}^{n}\left[RTS(t+1)\cdot e_k(t+1) + RTS(t)\cdot e_k(t)\right]\cdot(\ln x_k^{t+1} - \ln x_k^t) \\
& + \frac{1}{2}\sum_{j=1}^{m}\left[\left(R_j^{t+1} - \varepsilon_j(t+1)\right) + \left(R_j^t - \varepsilon_j(t)\right)\right]\cdot\left(\ln y_j^{t+1} - \ln y_j^t\right) \\
& + \frac{1}{2}\sum_{k=1}^{n}\left[\left(e_k(t+1) - S_k^{t+1}\right) + \left(e_k(t) - S_k^t\right)\right]\cdot\left(\ln x_k^{t+1} - \ln x_k^t\right)
\end{aligned}$$

Equation (13) decomposes the Törnqvist productivity index into five sources. The first three terms measure the pure productivity growth (i.e. changes in efficiency, technical change and the scale effect). The following two terms are related to output and input aggregation biases. Note that the Törnqvist productivity index (12) uses revenue and cost shares to aggregate respectively outputs and inputs. If observed shares are not equal to output and input distance function shares, the Törnqvist index will be a biased measure of the pure productivity growth. The output aggregation biases measures any effect that non-marginal cost pricing may have on productivity growth because firms are not allocative efficient with respect to output prices or firms are no price-takers in output markets. This term vanishes, however, under proportional mark-up pricing or if outputs change at the same rate. On the other side, the input aggregation biases occurs because the firms are allocative inefficient with respect to input prices. This term vanishes when firms are input allocative efficient, then $e_k(t)=S_k^t$, or inputs change at the same rate.

In summary, once a translog output distance function is estimated, equation (13) provides a decomposition of the Törnqvist productivity index into three general sources: technical change, scale economies and technical efficiency change. It also allows for evaluating output and input aggregation biases. It is noteworthy that this equation leads to an exact decomposition suitable for discrete data. That is, it can be implemented without having to approximate concepts that are specified in the continuous time framework, just like in Bauer (1990).

3. Empirical illustration: the Spanish savings banks

In this section, the parametric decomposition of the generalized Malmquist productivity index developed above is calculated using panel data from Spanish savings banks. This decomposition is based on the estimation of a translog output-oriented distance function using econometric techniques.

3.1. Econometric specification

Following Färe and Primont (1996), the econometric version of the output distance function can be written as:

$$\ln 1 = \ln D_o(y_{it}, x_{it}, t, \beta) + v_{it} + u_{it} \qquad (14)$$

where subscript "i" stands for firms, β is a vector of parameters to be estimated, u_i is a non-negative disturbance allowing for technical inefficiency, and the stochastic nature of the frontier is modeled by adding a two-sided random error term v_{it}. The following distribution assumptions are imposed:

$$\begin{aligned}
&v_{it} \rightarrow N(0, \sigma_v^2)\\
&u_{it} = u_i \cdot \exp\left[-\eta_1(t-T) - \eta_2(t-T)^2\right] \quad , \quad t \in \psi(i), \quad i = 1,\dots, N\\
&u_i \rightarrow N^T(0, \sigma_u^2) \quad , \quad u_i \geq 0
\end{aligned}$$

where $\eta=(\eta_1\ \eta_2)$ is a vector of parameters to be estimated, N is the total number of firms in the sample, $\psi(i)$ represent the set of T_i time periods among the T periods involved for which observations for the i*th* firm are obtained. That is, the noise term v_{it} is assumed to follow a normal distribution with mean zero and variance σ_v^2. The inefficiency term u_{it} is modeled as the product of an exponential function of time and a non-negative time-invariant firm effect, u_I. The term u_I is the inefficiency level of the *i*th producer at time T and is supposed to come from a non-negative truncated normal distribution with mean zero and variance σ_u^2.

Note that the exponential function of time is a generalization of the model by Battese and Coelli (1992). They assumed η_2=0 above, which is a rigid parameterization since efficiency must either increase at a decreasing rate (i.e. η_1>0), decrease at an increasing rate (i.e. η_1<0) or remain constant (i.e. η_1=0). In order to relax monotonicity, I propose a two-parameter specification, which allows for decreasing and increasing efficiency over the whole period considered.

The model defined by equation (14) can only be estimated as long as the homogeneity restrictions are imposed. The linear homogeneity in output quantities is imposed normalizing, using an arbitrary output as numeraire, the constant regressand and the other outputs.[8] Using the following notation $y^*_{jit}=y_{jit}/y_{mit}$ (j=1...m-1), the translog version of model (14) can be written as:

$$
\begin{aligned}
-\ln y_{mit} = \Bigg[& \alpha_0 + \sum_{k=1}^{n} \alpha_k \ln x_{kit} + \sum_{j=1}^{m-1} \beta_j \ln y^*_{jit} \\
& + \frac{1}{2}\sum_{k=1}^{n}\sum_{h=1}^{n} \alpha_{kh} \ln x_{kit} \ln x_{hit} + \frac{1}{2}\sum_{j=1}^{m-1}\sum_{h=1}^{m-1} \beta_{jh} \ln y^*_{jit} \ln y^*_{hit} \\
& + \sum_{k=1}^{n}\sum_{j=1}^{m-1} \gamma_{kj} \ln x_{kit} \ln y^*_{jit} \\
& + \psi_0 t + \frac{1}{2}\psi_{00} t^2 + \sum_{k=1}^{n} \xi_{kt}\, t \ln x^t_k + \sum_{j=1}^{m-1} \tau_{jt}\, t \ln y^*_{jit} \Bigg] + v_{it} + u_{it}
\end{aligned}
\tag{15}
$$

3.2. Data and sample

The data are yearly data (1985-98) from the Spanish Savings banks Confederation. The number of banks decreased throughout the sample period due to mergers and acquisitions. These mergers took place especially among savings banks. In this paper, a merged institution is treated as a different bank from the institution that existed before the merging process. That is, if two institutions merged into a new one they disappear and a new bank is born.

To select the variables I follow the majority of the literature and apply the intermediation approach proposed by Sealey and Lindley (1977) which treats deposits as inputs and loans as outputs. Four types of outputs and three types of inputs are included. The outputs are (OA) Bonds, cash and others assets not covered by the following outputs; (LO) Loans to no-banks; and (NI) noninterest income. Using this last output goes beyond the intermediation approach as commonly modeled. I include noninterest income in an attempt to capture off-balance-sheet activities such as securitization, brokerage services, and management of financial assets, which are becoming increasingly important at Spanish banks. The inputs are: (D) time and savings deposits; (F) deposits from banks and other funds; (L) Labor, measured by personnel expenses; and (K) Capital, measured by physical capital depreciation and other non-interest expenses. A summary of the descriptive statistics of these variables can be found in Table 1. All monetary variables are expressed in millions of pesetas and in real terms of 1985 by deflating by the GDP deflator index.

All variables have been mean-corrected prior to estimations. That is, each output and input variable has been divided by its geometric mean. In this way, the first order coefficient can be interpreted as distance elasticities evaluated at the sample mean. In addition, linear homogeneity in outputs is imposed using output OA as a numeraire.

8 This method is widely described in Coelli and Perelman (1996).

3.3. Empirical results

The output distance function (14) is estimated allowing error structure to differ between no-merged and merged banks. Since a merger process involve important structural changes (closure of branches, staff reallocation, etc.), the temporal pattern of technical efficiency of merged banks may be quite different to that of no-merged banks. Therefore, both groups of banks do not share likely the same random (efficiency) distributions.

The empirical results for the estimated model are presented in Table 2. All the elasticities have the expected signs at the geometric mean. Therefore, the estimated distance function fulfills at this point the property of monotonicity (i.e. non-decreasing in outputs and decreasing in inputs). The sum of the input elasticities is -1.0303 and is significantly different from zero, indicating the presence of moderate increasing returns to scale at the mean, as found in many past banks studies.

Following Battese and Coelli (1992) the estimated parameters can be used to calculate indexes of technical efficiency. Summary statistics of the technical efficiency indices appear in Table 3. The results shown in this table reveal that non-merged banks are on average more efficient than merged banks. Other feature of technical efficiency is noteworthy. Our model allows assessing the variations in technical efficiency over time. A likelihood ratio test is used to determine whether the data supported a model with time-invariant cost efficiency, i.e. imposing that $\eta_1=\eta_2=0$. While this hypothesis was rejected for merged banks, it was not rejected for the group of non-merged banks. This result indicates that the measured variations in efficiency for non-merged savings banks over the period 1985-98 are not statistically significant. This result conforms to the results obtained in the majority of empirical Spanish banking studies (see, for example, Maudos, 1996).

Total factor productivity growth can be measured from the estimated output distance function using equation (10). This combines the effect of technical change, technical efficiency changes and a scale effect as outputs expand over time. Table 4 reports the annual decomposition of growth of total factor productivity, together with an overall year average and three sub-period averages, for non-merged banks. The results show an increase of total factor productivity over the entire period. This is largely attributable to a strong technical progress and the positive, but modest, effect of scale economies. On the contrary, the efficiency changes have not had an important effect on productivity growth. Overall, these results are quite similar to those found by Pastor (1995), Maudos (1996) and Grifell and Lovell (1997), although their results rely on data up to 1994.

Table 5 reports the decomposition of growth of total factor productivity for merged banks. These banks experienced increased productivity since the early 1990s. This increase was especially intense since 1994. Through 1994 productivity growth can be attributed mainly to technical progress, whereas the efficiency effect is negative, but after 1990 the productivity increase is a consequence of improvements in both technical efficiency and technical progress. As expected, the scale effect in merged banks had a smaller effect on productivity growth than in non-merged banks.

4. Summary and conclusions

The primary objective of this paper is to decompose a generalized Malmquist-type productivity index into terms related to technical change, technical efficiency changes and returns to scale. This decomposition is based on parametric estimation of translog output-oriented distance functions and draws on the so-called exact index number approach to the derivation of productivity change measures. This approach also allows us decomposing the traditional Törnqvist productivity index without having to approximate concepts that are specified in the continuous time framework.

The parametric decomposition of the generalized Malmquist productivity index was applied to panel data coming from Spanish savings banks. The results show an increase of total factor productivity for both merged and non-merged banks. The main factor contributing to this increase was a strong technical progress. I also find that returns to scale have also a positive effect on productivity growth, indicating that the scale effect should be included in an examination of bank productivity growth.

References

Balk, B.M. (1999): "Scale Efficiency and Productivity Change", Mimeo.

Balk, B.M. (1993): "Malmquist Productivity Indexes and Fisher Ideal Indexes: Comment", *Economic Journal*, 103, 680-682.

Battese, G. and T. Coelli (1992): "Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India", *Journal of Productivity Analysis*, 3, 153-69.

Bauer, P. (1990): "Decomposing TFP Growth in the Presence of Cost Inefficiency, Nonconstant Returns of Scale, and Technological Progress", *Journal of Productivity Analysis*, 1, 287-301.

Caves, D.W., Christensen, L.R. and Diewert, W. (1982): "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity", *Econometrica*, 50(6), 1393-1414.

Coelli, T. and S. Perelman (1996): "Efficiency measurement, multiple-output technologies and Distance functions: With application to European Railways", CREPP Working Paper, 96/05. (University of Liege, Belgium).

Denny, M., Fuss, M. and Waverman, L. (1981): "The Measurement and Interpretation of Total Factor Productivity in Regulated Industries with an Application to Canadian Telecommunications", in Cowing, T. and Stevenson, R. (eds.), *Productivity Measurement in Regulated Industries*, New York, Academic Press, 172-218.

Diewert, W. (1976): "Exact and Superlative Index Numbers", *Journal of Econometrics*, 4, 115-145.

Färe, R. and Grosskopf, S. (1992): "Malmquist Productivity Indexes and Fisher Ideal Indexes", *Economic Journal*, 102, 158-160.

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, 85-102.

Färe, R., Grosskopf, S., Norris, M. and Zhang, Z. (1994): "Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries", *American Economic Review*, 84(1), 66-83.

Färe, R. and Primont, D. (1985): *Multi-Output Production and Duality: Theory and Applications*, Kluwer Academic Publishers, Boston.

Färe, R. and Primont, D. (1996): "The opportunity cost of Duality", *Journal of Productivity Analysis*, 7, 213-224.

Farrell, M.J. (1957): "The Measurement of Productive Efficiency", *Journal of the Royal Statistical Society*, 120, 253-281.

Fuentes, H., Grifell, E., and S. Perelman (1998): "A parametric distance function approach for Malmquist index estimation", Working Paper CREPP 98/03, Université de Liège.

Grifell, E. and Lovell, C.A.K. (1995): "A Note on the Malmquist Productivity Index", *Economic Letters*, 47, 169-175.

Grifell, E. and Lovell, C.A.K. (1997): "The sources of productivity change in Spanish banking", *European Journal of Operational Research*, 98, 364-380.

Grifell, E. and Lovell, C.A.K. (1999): “A Generalized Malmquist Productivity Index”, *Sociedad de Estadística e Investigación Operativa*, 7, 81-101.

Maudos, J. (1996): “Eficiencia, cambio tecnológico y productividad en el sector bancario español: una aproximación de frontera estocástica”, *Investigaciones Económicas*, 20(3), 339-358.

Pastor, J. (1995): "Eficiencia, cambio productivo y cambio tecnológico en los bancos y cajas de ahorro españolas: un análisis frontera no paramétrico", *Revista de Economía Española*, 12(1), 35-73.

Sealey, C.W. and Lindley, J.T. (1977): “Inputs, Outputs, and a Theory of Production and Cost at Depository Financial Institutions”, *Journal of Finance*, 32(4), 1251-1266

Table 1
Descriptive statistics of selected variables

Variable	Mean	St. Dev.	Minimum	Maximum
OA	74068	167019	583	1737633
LO	134773	248454	1066	2641366
NI	1473	3863	5	51670
D	225808	409165	2802	3539309
F	31978	95042	38	1184926
L	4515	7423	69	68081
K	2923	5103	33	46767

Table 2
Estimated parameters of equation (15)

Variable	Coefficient	Value	t-test	Variable	Coefficient	Value	t-test
Constant	α_0	0.0716	3.296	Ln(D)· Ln(LO)	γ_{12}	0.0657	0.882
Ln(NI)	β_3	0.0445	1.853	Ln(F)· Ln(LO)	γ_{22}	-0.0135	-1.004
Ln(LO)	β_2	0.6800	19.161	Ln(L)· Ln(LO)	γ_{32}	-0.0639	-0.973
Ln(D)	α_1	-0.6561	-11.133	Ln(K)· Ln(LO)	γ_{42}	0.0106	0.171
Ln(F)	α_2	-0.0634	-5.138	Ln(D)· Ln(F)	α_{12}	0.0019	0.058
Ln(L)	α_3	-0.1254	-2.368	Ln(D)· Ln(L)	α_{13}	-0.3537	-2.329
Ln(K)	α_4	-0.1774	-2.602	Ln(D)· Ln(K)	α_{14}	0.0483	0.373
$^1/_2$Ln(NI)^2	β_{33}	-0.0688	-2.425	Ln(F)· Ln(L)	α_{23}	0.0254	0.934
$^1/_2$Ln(LO)^2	β_{22}	0.0236	0.529	Ln(F)· Ln(K)	α_{24}	0.0220	0.830
$^1/_2$Ln(D)^2	α_{11}	0.3028	2.462	Ln(L)· Ln(K)	α_{34}	-0.0292	-0.301
$^1/_2$Ln(F)^2	α_{22}	-0.0455	-5.503	t	ψ_0	-0.1383	-3.970
$^1/_2$Ln(L)^2	α_{33}	0.3271	1.840	$^1/_2$ t^2	ψ_{00}	-0.0052	-0.201
$^1/_2$Ln(K)^2	α_{44}	-0.0208	-0.139	t· Ln(NI)	τ_3	0.0352	1.815
Ln(NI)· Ln(LO)	α_{23}	0.0964	3.430	t· Ln(LO)	τ_2	-0.0092	-0.339
Ln(D)· Ln(NI)	γ_{13}	0.0746	1.397	t· Ln(D)	ξ_1	-0.0229	-0.471
Ln(F)· Ln(NI)	γ_{23}	-0.0056	-0.548	t· Ln(F)	ξ_2	0.0160	1.935
Ln(L)· Ln(NI)	γ_{33}	-0.0393	-0.665	t· Ln(L)	ξ_3	-0.0009	-0.023
Ln(K)· Ln(NI)	γ_{43}	-0.0174	-0.260	t· Ln(K)	ξ_4	0.0016	0.029
Non merged banks				Merged banks			
	σ^2	0.0188	4.243		σ^2	0.0199	2.695
	λ	0.2593	3.303		λ	0.1177	2.211
	η_1	0.0192	0.612		η_1	0.2158	4.749
	η_2	0.0014	0.575		η_2	0.0309	5.844

Table 3
Descriptive statistics of technical efficiency indexes (%)

	Non merged banks			Merged banks		
Year	Mean	St. Dev.	Number	Mean	St. Dev.	Number
1985	91.1	6.57	77	.	.	.
1986	91.0	6.66	77	.	.	.
1987	90.9	6.73	77	.	.	.
1988	90.8	6.79	77	.	.	.
1989	90.9	6.86	75	93.9	0	1
1990	91.0	5.89	54	91.4	6.30	10
1991	90.8	6.22	42	89.1	7.89	13
1992	91.1	6.16	39	86.1	8.06	14
1993	90.9	6.36	36	84.7	8.74	15
1994	91.0	6.32	36	83.9	9.18	15
1995	91.0	6.42	34	84.1	8.92	16
1996	91.1	6.35	34	84.9	8.50	16
1997	91.2	6.27	34	86.5	7.70	16
1998	91.4	6.17	34	88.6	6.60	16
1985-1998	91.0	6.44	726	86.4	8.17	132

Table 4. Total factor productivity growth. Non merged banks

Period	Productivity growth	Changes in Efficiency	Technical Change	Scale Effect
85/86	2.84	-0.15	2.87	0.12
86/87	2.97	-0.13	2.88	0.22
87/88	3.08	-0.10	2.86	0.32
88/89	3.00	-0.07	2.81	0.27
89/90	2.94	-0.04	2.89	0.09
90/91	3.02	-0.02	2.91	0.12
91/92	2.84	0.01	2.69	0.14
92/93	2.69	0.04	2.51	0.14
93/94	2.70	0.06	2.48	0.16
94/95	2.75	0.09	2.51	0.15
95/96	2.85	0.12	2.56	0.17
96/97	2.87	0.14	2.53	0.19
97/98	2.75	0.16	2.44	0.15
85/90	2.97	-0.10	2.86	0.21
90/94	2.82	0.02	2.66	0.14
94/98	2.80	0.13	2.51	0.17
85/98	2.90	-0.03	2.74	0.18

Note: Arithmetic average of annual rates in percentages.

Table 5. Total factor productivity growth. Merged banks

Period	Productivity growth	Changes in Efficiency	Technical Change	Scale Effect
85/86	.	.	.	.
86/87	.	.	.	.
87/88	.	.	.	.
88/89	.	.	.	.
89/90	0.73	-2.26	2.81	0.18
90/91	0.33	-2.60	2.83	0.10
91/92	0.10	-2.79	2.72	0.17
92/93	0.58	-2.11	2.61	0.08
93/94	1.70	-1.10	2.64	0.16
94/95	2.73	-0.01	2.66	0.09
95/96	3.87	1.06	2.68	0.13
96/97	4.78	1.94	2.68	0.16
97/98	5.27	2.50	2.61	0.16
85/90	.	.	.	.
90/94	0.77	-2.04	2.68	0.13
94/98	4.19	1.40	2.66	0.14
85/98	2.69	-0.11	2.67	0.13

Note: Arithmetic average of annual rates in percentages.