Knowledge Production Function and Malmquist Index Regression Equations as a Dynamic System

JINGHAI ZHENG

Department of Economics
University of Gothenburg, Sweden

Paper submitted to a Special Issue in Memory of Lennart Hjalmarsson
with Journal of Productivity Analysis

June 2012

Abstract
When considering the linkage between innovation and productivity, the relationship is often termed as the knowledge production function in the endogenous growth literature. In this article, we approach the issues involved through the aspects of the standard neoclassical production theory. The advantage is that the basic properties required of ordinary production function can be employed to infer the microstructure of the knowledge acquisition process. This approach is also attractive in an empirical sense that the techniques of applied productivity analysis might be effective in investigations on the determinants of productivity and economic growth. As an example, we demonstrate that when the popular DEA-based Malmquist productivity indexes are used in regression analysis the set of linear equations involved can be treated as a system. With reference to the special structure of the knowledge production function, the regression equations can be further specified as a dynamic system. The properties of the Malmquist Index regression equations provide rich microstructures for the relationship between productivity growth, productivity growth components, and their determinants.

JEL classification: L10; L96; O30

---

1 Jinghai Zheng is also Guest Research Fellow at the Centre for China Studies, Tsinghua University, Beijing, China.
1. Introduction

Simar and Wilson (2007) note that searching results on internet in 2004 showed that more than 800 items found associated with working papers using the two-stage procedure for analysis on determinants of DEA scores, in addition to the more than 40 published papers included in their survey of the literature. The two-stage procedure in fact has been well received in the literature on both theoretical and statistical grounds. Ray (1991) gave two reasons — one theoretical and one practical — for using the procedure: First, researchers can treat the non-production factors as fixed and “indivisible non-discretionary” inputs and do not need the free disposability assumption with respect to these inputs; Second, it gives researchers substantial flexibility to choose non-production inputs in regression analyses. Simar and Wilson noted that several authors including Chilingerian and Sherman (2004), Ray (2004), and Ruggiero (2004) acknowledged the usefulness of the two-stage procedure in the analysis of the technical efficiency and productivity determinants.

The two-stage procedure also possesses some favorable statistical properties. Banker (1993) discusses the statistical basis for the DEA efficient scores and demonstrates that DEA estimators of best practice are maximum likelihood estimators under rather general assumptions and “exhibit the desirable asymptotic property of consistency”. Banker and Natarajan (2008) further show that “the two-stage procedure consisting of DEA followed by ordinary least squares (OLS) regression analysis yields consistent estimators of the impact of contextual variables”. In other words, the two-stage procedure can be seen as a stochastic production frontier with “both one-sided inefficiency deviations as well as two-sided random noise”. Their study thus provides a formal statistical basis for regressing DEA efficiency scores (perhaps as well as Malmquist productivity indexes) on explanatory variables.

For analysis using small samples, advances in the bootstrapping DEA scores (Simar and Wilson 1998, 2000a, and 2000b) and Malmquist indexes (Simar and Wilson 1999) have further made the non-parametric estimation of technical efficiency and productivity growth more attractive among researchers. As to the two-stage procedure, Simar and Wilson (2007) have resolved a key problem with the second step regression analysis. Xue and Harker (1999) noted that “DEA efficiency scores are dependent on each other in the statistical sense” and the dependency was due to that DEA scores were not absolute efficiency measures. Similar concern can be raised about the dependency of the efficiency change and technical progress indexes with the Malmquist productivity index decomposition since these indexes are calculated on the basis of the DEA efficiency scores. Simar and Wilson (2007) took care of the problem by using a double bootstrap technique.
While studies using the two step procedure for analyzing DEA efficiency scores have been plenty such as those found in Destefanis and Sena (2007), Madden and Savage (2001), and Zheng, Liu, and Bigsten (1998), there have been much fewer applications in which the methodology was applied to the Malmquist productivity index and its components. Examples using Malmquist indexes as explanatory variables in regression analysis include, Castellacci and Zheng (2008), Zheng, Liu, and Bigsten (2003), and Dismuke and Sena (1999).

Since the early 1980s economists have started to decompose total factor productivity growth into technical progress and technical efficiency change components (Nishimizu and Page 1982). In recent years the DEA-based Malmquist index model of productivity decomposition (Färe et al 1994) appears to have become standard in the investigation on productivity performance of firms and regions. A major advantage of this approach is that it allows decomposing the change in total factor productivity (TFP) into technical progress and technical efficiency change; loosely speaking, a shift in the best-practice production frontier is an indication of technical change, and technical efficiency change is associated with learning by doing, improved managerial practices, and change in the efficiency when using an existing technology (Felipe, 1999). Once TFP growth is decomposed into efficiency change and technical progress, we can identify the explanatory variables affecting a specific component. By separating the best practice firms from inefficient ones, we can evaluate the performance of the firms with different characteristics.

Combining the Malmquist index estimates with regression analyses forms a multi-equation system for examination of the links between TFP indexes (in both level and growth rate) and their determinants in a two-stage procedure (for example in Zheng, Liu, and Bigsten 2003). The advantages of such regressions are obvious since multiple equations would allow researchers to identify the channels through which a specific factor affects TFP growth if different factors are behind different productivity growth components. Supplemented with technical efficiency (level) index and a binary variable representing best practice firms, the relationships between the regression equations can become even more complicated. Therefore, the decomposition of TFP and the possibility of interactions between TFP level and growth rate, and between technical progress and technical efficiency change components present both challenges and opportunities for empirical analyses.

The basic idea of this study is to propose in a formal manner a production function framework that can be established when considering the relationship between innovation and productivity. The relationship has been referred to as “knowledge production function” in

---

2 Destefanis and Sena (2007) used DEA technical efficiency scores and a binary variable to investigate their determinants with the OLS regression and estimation of a Logit model.

3 Dismuke and Sena (1999) regressed Malmquist productivity indexes on a set of explanatory variables using SURE estimation methods without explicit knowledge about the structure of such a system.
Romer (1990), but we proceed with a formal presentation that is similar to the “technology function” in Phelps (1966).

The order of our presentation is as follows. Section 2 introduces the basic assumptions concerning the structure of the knowledge production function that is consistent with the Malmquist indexes as measures of knowledge flows. Conditions required of the Malmquist indexes to be treated as measures of the knowledge flows are presented in Section 3. Taking account of the special structure for the knowledge production function, in Section 4 we demonstrate that the Malmquist index regression equations can be regarded a dynamic system with important theoretical implications and useful properties for empirical analysis. Section 5 provides an empirical example using Norwegian firm data and we summarize in Section 6.

2. Basic Assumptions

Although the DEA-based Malmquist indexes have become increasingly popular in empirical works, so far it is not clear how regression analysis should be carried out if all the indexes are used simultaneously regressing on a set of explanatory variables. In what follows we demonstrate that given the structure of the Malmquist index model, the popular DEA-based Malmquist indexes when used in regression analysis the set of linear equations involved can be treated as a dynamic system.

Investigation on the determinants of productivity is to endogenize total factor productivity (TFP) either in level or growth rate or both with respect to factors that may impact productivity performance. In endogenous growth literature this relationship is often referred to as the knowledge production function as in Romer (1990). However, specifications involved so far in the literature are often simplistic and restrictive as they usually take the form of Cobb-Douglas. Several examples can be found in Jones (1999). Since any specification for the relationship between TFP and its determinants can be seen as a knowledge production function, we would like to see what major assumptions are involved at the basic level.

When the production unit is the firm, its production function may be given by the standard Hicks-neutral representation:

\[ y = AF(x) \] (2.1)

where \( y \) is output and \( x \) are inputs. To be more specific, we write it more specifically as,

\[ y = AF(K, L_t) \] (2.1B)
where \( Y \) is the ordinary output, \( K \) and \( L_Y \) stand for capital and labor, and \( A \) refers to the level of technology. The knowledge production is measured as the change in knowledge stock \( \dot{A} \) (Romer 1990), so we write the basic knowledge production function as follows.\(^4\)

\[
\dot{A} = f(x) \tag{2.2}
\]
or more specifically as in Phelps (1966),

\[
\dot{A} = f(K_A, L_A, L, A) \tag{2.2B}
\]

where \( K_A \) and \( L_A \) are technology capital and research labor input. Note that \( L = L_A + L_Y \) is the labor force of the firm. The knowledge production function may maintain the following properties:

A1. The knowledge production function is monotone in inputs
A2. The input set of the knowledge production is convex
A3. Researcher is the essential input, i.e., \( \dot{A} = f(K_A, 0, L, A) = 0 \)
A4. The input set is nonempty and closed
A5. The knowledge production function is finite, nonnegative, real valued, and single valued for all nonnegative and finite inputs

and the knowledge production function \( f(\cdot) \) is the derivative of the “technology function” of Phelps (1966):

\[
A(t) = \int_{-\infty}^{t} f(E(L_A(v), L(v), K_A(v)), A(v - \omega))dv \tag{2.3}
\]

Phelps (1966) gave a list of the dynamic properties for a valid “technology function”, \( A(t) \) as follows:

B1. Diminishing returns
B2. Diminishing marginal rate of substitution
B3. The marginal effectiveness of current research is an increasing function of the level of technology recently attained (technical progress in research)
B4. Exponential growth of researchers will produce an exponential increase of the level of technology

\(^4\) Phelps (1966) uses the term “technology function” in referring to the relationship between research inputs and research outcomes; Gumolka (1970) mentioned “the production function of innovations”; and Jones (1999, 2002, and 2006) prefers to phrase the relationship as “the idea production function.”
Due to the dynamic nature of the knowledge production function not all specifications are feasible for theoretical analyses and empirical estimations. Therefore, one possible specification is a Hicks-neutral form as follows:

\[ \dot{A} = f(K_A, L_A, A) = \Lambda \phi(K_A, L_A) \]  

(2.4)

Thus the knowledge level is multiplicatively separable from the rest of the function. Phelps (1966) referred to \( \phi(K_A, L_A) \) as the “effective research function”. Note that the formulation (2.4) is not usually homogeneous of degree one in its arguments unless \( \phi(K_A, L_A) \) is homogeneous of degree zero (which is unlikely). This property of (2.4) will have consequences in the dynamic setting.

Like in Nordhaus (1969), to secure linear homogeneity of the entire knowledge production function it is necessary to consider a more flexible form as in Jones (1999) as follows:

\[ \dot{A} = A^\phi \phi(K_A, L_A) \]  

(2.5)

where \( \phi < 1 \) is imposed. This specification presents “increasing returns in the production of new ideas” when \( \phi > 0 \). It also allows for diminishing returns if \( \phi < 0 \). Its Solow residual form (Solow 1957) is given by

\[ \dot{A} = A^{\phi-1} \phi(K_A, L_A) \]  

(2.6)

Log linearization of (2.6) yields:

\[ \ln A = (\phi-1)A + \ln \phi(K_A, L_A) \]  

(2.7)

This result is interesting in that knowledge stock enters the relationship linearly and \( \ln \phi(K_A, L_A) \) can be specified with a log linear function such as Cobb-Douglas and translog, which makes the empirical formulation easy on one hand. For example in the Cobb-Douglas case we will have:

\[ \ln A = \delta + (\phi-1)\ln A + \alpha \ln K_A + \beta \ln L_A \]  

(2.8)

On the other hand the appearance of Solow residual in log, \( \ln \frac{\dot{A}}{A} \) presents some difficulties in empirical work because in practice Solow residual can be negative.

---

5 Concerning the “the future of the new economy”, Jones (2001) notes “Economists cannot say what it takes to generate knowledge at a permanently faster rate and thereby raise the productivity growth rate permanently” and “It is certainly possible that the economy becomes increasingly better at producing new ideas, … However, it is also possible that it becomes increasingly difficult to discover new ideas, as the most obvious ideas are discovered first.”
3. Malmquist Productivity Indexes as Measures of Knowledge Flow

Economists have long known that TFP growth can be interpreted as technical progress under certain assumptions (Chambers, 1988). This interpretation was typically adopted in Solow (1957) for the well-known Solow residual, while the concept of technical efficiency was developed independently in Farrell (1957). Nishimizu and Page (1982) demonstrated that TFP growth or the Solow residual could be partly interpreted as technical progress and partly as change in technical efficiency since the definitions of the two concepts are mutually exclusive. Technical progress is defined as the shift of the production frontier over time while technical efficiency change refers to the movement towards the production frontier. However, more meaningful interpretations are required in the situation when TFP growth is regarded as endogenous.

For example, R&D expenditure alone will not be able to shift production frontiers or improve technical efficiency. It has to be used to add new expertise to the knowledge stock of the firm, and the new knowledge has to be used efficiently during which process additional knowledge (experience) are also accumulated. In the endogenous growth literature it is well accepted that the knowledge stock in turn can also be used to generate new knowledge, which is interpreted as the “standing on the shoulder of the giant” effect (Romer 1990).

We next show that the popular Malmquist TFP growth index can not only be decomposed into technical progress and technical efficiency change components as in Färe et al. (1994), but it can also be used as measure of knowledge flow that is consistent with the interpretation of TFP in the endogenous growth literature using an important result from Färe and Grosskopf (1996). Another important and useful result we would like to show is that although the Malmquist productivity index components are derived as multiplicatively separable, it is also approximately additively separable if a simple linearization scheme is applied.

Following Färe et al. (1994) the output-based Malmquist index of productivity change can be given by:

\[ S_t = \{ (x^t, y^t) : x^t \text{ can produce } y^t \} \]  
(3.1)

where time is represented by \( t = 1, \ldots, T \), the production possibility set \( S_t \) contains inputs, \( x^t \in \mathbb{R}^N \) and outputs, \( y^t \in \mathbb{R}^M \). Färe et al. also defines the output distance function at \( t \) as:

\[ D_o^t (x^t, y^t) = \inf \{ \lambda : (x^t, y^t/\lambda) \in S_t^t \} = (\sup \{ \lambda : (x^t, \lambda y^t) \in S_t^t \})^{-1} \]  
(3.2)

They noted two important properties. The first is that \( D_o^t (x^t, y^t) \leq 1 \) is both necessary and sufficient for \( (x^t, y^t) \in S_t^t \). The second property is that \( D_o^t (x^t, y^t) = 1 \) is both necessary and sufficient for \( (x^t, y^t) \) being on the production frontier.

When production technology exhibits constant returns to scale, the output distance function is homogenous of degree 1 in output and -1 in inputs (Färe and Grosskopf, 1996). The distance function can involve two different time periods as follows:

\[ D_o^t (x^{t+1}, y^{t+1}) = \inf \{ \lambda : (x^{t+1}, y^{t+1}/\lambda) \in S_t^t \} \]  
(3.3)
Färe et al., assuming constant returns to scale, specify the output-based Malmquist productivity change index as follows:

\[
M_o(x^{t+1}, y^{t+1}, x', y') = \left[ \frac{D'_o(x^{t+1}, y^{t+1})}{D'_o(x', y')} \left( \frac{D_o^{t+1}(x^{t+1}, y^{t+1})}{D_o^{t+1}(x', y')} \right) \right]^{1/2}.
\] (3.4)

They showed that the Malmquist productivity index in (3.4) could be further decomposed into two components:

\[
\text{EFFCH} = \frac{D'_o(x^{t+1}, y^{t+1})}{D'_o(x', y')},
\] (3.5)

and

\[
\text{TECH} = \left( \frac{D'_o(x^{t+1}, y^{t+1})}{D'_o(x^{t+1}, y^{t+1})} \frac{D'_o(x', y')}{D'_o(x', y')} \right)^{1/2},
\] (3.6)

where the expression in (3.5) measures the change in efficiency. Expression (3.6) indicates shifts in the frontier technology.

Another important property with the Malmquist index defined in (3.4) is that it is exactly the straight measure of productivity change when there are only one input and one output. For example, to show this Färe and Grosskopf (1996) defined TFP change in the one-input and one-output case as follows:

\[
\text{TFPCH} = \frac{y^{t+1}}{x^{t+1}} \frac{y'}{x'}
\]

Recall that output distance function is homogeneous of degree 1 in output and –1 in input under constant returns to scale, 

\[
D'_o(x', y') = (y'/x')D'_o(1,1)
\]

This leads to

\[
\text{TFPCH} = \frac{y^{t+1}}{x^{t+1}} \frac{y'}{x'} = \frac{y^{t+1}}{x^{t+1}} \frac{D'_o(1, 1)}{D'_o(1, 1)} = \frac{D'_o(x^{t+1}, x^{t+1})}{D'_o(x', y')}
\] (3.7)

which explains the first term in (3.4) as a productivity index. Subtracting 1 from this productivity index, it coincides with the TFP growth rate or Solow residual, \(\dot{A}/A\) in (3.4), which was interpreted in Romer (1990) as growth rate in the stock of knowledge.

A more formal proof on the equivalence of the Malmquist index and the usual Solow residual can be found in Fried, Knox, and Schmidt (2008) as follows. Assuming one output, production function is given by

\[
F(x) = \max(y : (x, y) \in S)
\]

Using constant returns to scale, this production function can be related to the output distance function:
\[ F(x) = y : D_o(x, y) = y / \left[ yD_o(x, 1) \right] = 1 / D_o(x, 1) \]

or

\[ D_o(x, y) = y / F(x) \]

If technical change is Hicks-neutral as in Solow (1957), a discrete version is given by

\[ y' = F'(x') = A(t)\varphi(x'), \quad t = 0, 1 \]
\[ D'_o(x', y') = y'/A(t)\varphi(x'), \quad t = 0, 1 \]

Substitute the terms in the Malmquist productivity index (3.4) with these expressions:

\[
M_o(x^1, y^1, x^0, y^0) = \left[ \left( \frac{y'/A(0)\varphi(x')}{{y^0}/A(0)\varphi(x^0)} \right) \left( \frac{y'/A(1)\varphi(x')}{{y^0}/A(1)\varphi(x^0)} \right) \right]^{1/2}
\]

\[ = A(1)/A(0) \]

Thus the Malmquist productivity index also qualifies as a unit free measure of knowledge flow upon which a knowledge production function can be defined, assuming constant returns to scale, Hicks-neutrality, and technical efficiency. In general:

\[
\begin{align*}
\text{TFPCH} & \approx (A/A) + 1 \\
\text{EFFCH} & \approx (\dot{A}_e/A_e) + 1 \\
\text{TECH} & \approx (\dot{A}_r/A_r) + 1
\end{align*}
\]

(3.8)

and note that TFPCH=EFFCH·TECH, therefore define,

\[
[(A/A) + 1] = [(\dot{A}_e/A_e) + 1] \cdot [(\dot{A}_r/A_r) + 1]
\]

(3.9)

The following results will be very useful in both theoretical analysis and empirical work:
\[
\frac{A + \dot{A}}{A} = (A_r + A_r') \left( \frac{A_E + A_E'}{A_E} \right)
\]

\[
= \frac{(A_r + A_r')(A_E + A_E')}{A_r A_E}
\]

\[
= \frac{(A_r A_E + A_E \dot{A}_r + A_r \dot{A}_E + A_r A_E')}{A_r A_E}
\]

(3.10)

Define \( A = A_r A_E \), then we have

\[
\dot{A} = A_E \dot{A}_r + A_r \dot{A}_E + A_r A_E'
\]

(3.11)

Divide both sides of the equation by \( A \)

\[
\frac{\dot{A}}{A} = \frac{A_E \dot{A}_r + A_r \dot{A}_E}{A_r} + \frac{A_r A_E'}{A_E}
\]

(3.12)

Ignore the last second-order term, the Malmquist index based TFP growth rate is given by

\[
\frac{\dot{A}}{A} \approx \frac{\dot{A}_r}{A_r} + \frac{\dot{A}_E}{A_E}
\]

(3.13)

given that \( A = A_r A_E \). The importance of the result lies in that the multiplicatively separable Malmquist productivity indexes is approximately additively separable in terms of productivity growth rates. Take one more step,

\[
\frac{\dot{\overline{A}}}{A} + 1 \approx \left[ \frac{\dot{A}_r}{A_r} + 1 \right] + \left[ \frac{\dot{A}_E}{A_E} + 1 \right] - 1
\]

(3.14)

i.e.,

\[
\text{TFPCH} \approx \text{TECH} + \text{EFFCH} - 1
\]

(3.15)

In fact the multiplicatively separable Malmquist productivity index decomposition relationship can be linearized directly using Talor series expansion around (1,1) as follows:

\[
\text{TFPCH} = \text{TECH} \cdot \text{EFFCH}
\]

\[
\approx 1 + \text{TECH} - 1 + \text{EFFCH} - 1
\]

\[
= \text{TECH} + \text{EFFCH} - 1
\]

(3.16)

One gets the same result as in (3.15).
Strictly speaking, the knowledge production measured, say, with number of patents is only the supply side while the demand side decides whether the knowledge are useful or not. Productivity outcomes are the results of balancing supply of with the demand for useful knowledge.

4. The Malmquist Index Regression Equations as a Dynamic System

Many studies have estimated the Malmquist index and its components, but it was not clear how analysis of determinants of these indexes should be carried out. In some cases, only level variables such as technical efficiency were used in regression analysis perhaps due to the lack of structural guidance with respect to the regressions using the Malmquist productivity growth index and its components. For example, one issue encountered in Zheng, Liu, and Bigsten (2003) was whether the first-differences instead of the levels of the explanatory variables should be involved in the regression since Malmquist productivity growth indexes are intrinsically first-differenced indicators. Fortunately, the knowledge production function defined in Section 2 provides convenient theoretical guidance in this regard. In fact the framework of the knowledge production function can be very useful in many other important aspects. We now propose a straightforward approach for conducting analysis in this context, which is to sort out the properties of the Malmquist indexes that have been largely ignored in the existing literature. Incorporation of the dynamic aspects of the knowledge production function in the Malmquist index regression equations generates desirable model properties and brings new opportunities for detailed analysis on the determinants of the productivity performance. We will examine these useful properties by means of propositions. Proves are also provided and discussed whenever necessary.

The endogenous growth literature basically says nothing about productivity decomposition, except in Nelson and Phelps (1966) in a somewhat different context. The linear nature of the Nelson and Phelps specification can be treated as a system if one is to exploit the cross-equation relations in econometric estimations. This was to take advantage of the cross-equation restrictions in the Malmquist index regressions. It is now clear that the multiplicatively separable Malmquist index decomposition can be conveniently turned into additively separable indexes as we have already shown in (3.15) and a formal treatment of the issue will be provided in this section later.

Given the dynamic nature of the knowledge production function, endogenizing simultaneously multiple productivity indicators such as the Malmquist index estimates requires working with differential equation systems. In what follows, we will demonstrate that specifications of the system reveal relationships that were not so clearly seen when the Malmquist index regression equations are not treated as a system. For example, since knowledge stock is considered as a productive input in the process of acquiring new knowledge, does there exist a division of labor between different knowledge? In the context
of productivity growth decomposition, the question is whether knowledge accumulated through technical efficiency improvement will be useful when the objective is to produce a new product? On the other hand, we ask if one needs knowledge on past technologies when working to improve technical efficiency with newly installed machineries. If one is to decide the specification of the dynamic system, it is inevitable that these considerations need to be taken into account. A fully specified dynamic system based on the Malmquist index regression equations will provide new opportunities for in depth studies of the relationship between total factor productivity and its determinants.

To begin with, we present a “full” list of the Malmquist index regression equations as follows:

(1) Technical efficiency

\[ A = D_t'(x, y) \cdot \int_0^t f(x, A_t(v), A_e(v)) dv \]

(2) Best practice (binary)

\[ A_{\text{front}} = \int_0^t f(x, A_t(v), A_e(v)) dv \]

(3) Total factor productivity growth

\[ \frac{\dot{A}}{A} + 1 = 1 + f(x, A_t, A_e) \]

(4) Technical efficiency change

\[ \frac{\dot{A}_t}{A_t} + 1 = 1 + f_t(x, A_t, A_e) \]

(5) Technical progress

\[ \frac{\dot{A}_e}{A_e} + 1 = 1 + f_e(x, A_t, A_e) \]

where \( A_t \) stands for the level of technology and \( A_e \) the level of technical efficiency. When treating the equations (1) through (5) as a dynamic system, the following properties hold.

**Proposition 1** The system of linear regression equations on the basis of (4.4) - (4.6) is not independent.
This is obvious since multiplication of (4.5) and (4.6) is equal to (4.4). Estimation of any two equation enables researcher recover the third equation. Assume that

\[
\begin{align*}
\dot{A}_T/A_T + 1 &= \sum \beta_{T_i} x_i \\
\dot{A}_E/A_E + 1 &= \sum \beta_{E_i} x_i
\end{align*}
\]

where \( x \)'s are explanatory variables and \( \beta \)'s parameters to be estimated. The matrix form is given by,

\[
\begin{bmatrix}
\dot{A}_T/A_T + 1 \\
\dot{A}_E/A_E + 1
\end{bmatrix} =
\begin{bmatrix}
\beta_{T_1} & \beta_{T_2} & \ldots & \beta_{T_n} \\
\beta_{E_1} & \beta_{E_2} & \ldots & \beta_{E_n}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix}
\]

and note that the relationship between productivity growth and its components is

\[
[(\dot{A}/A) + 1] = [(\dot{A}_T/A_T) + 1] \cdot [(\dot{A}_E/A_E) + 1]
\]

\[
= \left( \sum \beta_{T_i} x_i \right) \cdot \left( \sum \beta_{E_i} x_i \right)
\]

\[
= \sum \beta_{T_i} \beta_{E_i} x_i x_j
\]

Collecting the equations as a system, we have

\[
\begin{align*}
[(\dot{A}/A) + 1] &= \sum \beta_{T_i} \beta_{E_i} x_i x_j \quad \text{(P1.1)} \\
[(\dot{A}_T/A_T) + 1] &= \sum \beta_{T_i} x_i \quad \text{(P1.2)} \\
[(\dot{A}_E/A_E) + 1] &= \sum \beta_{E_i} x_i \quad \text{(P1.3)}
\end{align*}
\]

which is a system of equations with cross-equation restrictions. This implies that parameters in (P1.1) can be derived from parameter estimates of (P1.2) and (P1.3), which means that only two equations need to be estimated.

**Proposition 2** If the firm is already on the production frontier then the only way for it to raise productivity is through technical progress.

The proof is rather straightforward. If a firm is already on the production frontier at time \( t \), then the best it can do at time \( t+1 \) is to be on the production frontier in order to avoid negative contributions to productivity growth in terms of technical efficiency change defined in (4.5).
Second, if technology does not improve at \( t+1 \), the second term in (4.6) is one, and the first term will also be one since frontier does not shift. Combining the two cases, equation (4.4) will be equal to one, indicating zero productivity growth (a contradiction).

This property is rather useful in the context of multiple regressions involving several productivity indexes defined here. Factors that always manage to push the firms to the frontier will not be effective in the regression for technical efficiency change. The efficiency change regression equation becomes redundant.

**Proposition 3** If the stock of a firm’s knowledge is itself a factor of production, then the Malmquist Index regression equations can be treated as a dynamic system.

**Proof:**
Denoting the Malmquist total factor productivity (TFP) index \( M_o \) with the notation in Jones (1999) as

\[
M_o \approx \frac{A + \dot{A}}{A} = \frac{\dot{A}}{A} + 1 \tag{P3.1}
\]

Since \( M_o \) can be further decomposed into \( M_T \) and \( M_E \), which are mutually exclusive, we have

\[
M_0 = M_T \cdot M_E \approx \frac{A_T + \dot{A}_T}{A_T} \cdot \frac{A_E + \dot{A}_E}{A_E} = \left( \frac{\dot{A}_T}{A_T} + 1 \right) \left( \frac{\dot{A}_E}{A_E} + 1 \right) \tag{P3.2}
\]

The specification for the determinants of both \( M_T \) and \( M_E \) are complicated. One possible choice is as follows:

\[
\begin{align*}
(M_T - 1) &\approx \frac{\dot{A}_T}{A_T} = A_T^{(\phi - 1)} f_T(x) \\
(M_E - 1) &\approx \frac{\dot{A}_E}{A_E} = A_E^{(\rho - 1)} f_E(x)
\end{align*} \tag{P3.3}
\]

which is a Hicks-neutral case for simplicity and it says that the rate of technical progress depends on the stock of technology itself, but not on the stock of knowledge associated with technical efficiency improvement and vice versa. Alternatively,

\[
\begin{align*}
(M_T - 1) &\approx \frac{\dot{A}_T}{A_T} = A_T^{(\phi - 1)} f_T(x) \\
(M_E - 1) &\approx \frac{\dot{A}_E}{A_E} = A_E^{(\rho - 1)} f_E(x)
\end{align*} \tag{P3.4}
\]
which specifies the entire knowledge stock of the firms can be used for new knowledge acquisition, i.e., the knowledge on technology itself and knowledge about how to improve efficiency. Zheng (2008) adopted a specification like (P3.4). Since this specification could be a result of casual inclusion of explanatory variables when conducting regression analysis on the estimates of the Malmquist indexes we will sometimes refer to it as the “naïve” regression below. Notice that \( A = A_T \cdot A_E \), hence a more flexible specification might be given by

\[
\begin{align*}
(M_T - 1) & \approx \frac{\dot{A}_T}{A_T} = A_T^{(\phi_T - 1)} A_E^{\phi_E} f_T(x) \\
(M_E - 1) & \approx \frac{\dot{A}_E}{A_E} = A_T^{\phi_E} A_E^{(\phi_E - 1)} f_E(x)
\end{align*}
\]

and \( \phi \neq \theta \) in general. When \( \phi = \theta \) the system becomes (P3.4), and it is reduced to (P3.3) with \( \phi = \theta = 0 \). We take the commonly encountered specification of (P3.4) in empirical applications as an example:

\[
M_0 \approx \left[ \frac{\dot{A}}{A} + 1 \right] = \left[ \frac{\dot{A}_T}{A_T} + 1 \right] \left[ \frac{\dot{A}_E}{A_E} + 1 \right] = \left( A_T^{(\phi_T - 1)} f_T(x) + 1 \right) \left( A_E^{(\phi_E - 1)} f_E(x) + 1 \right) 
\]

\[
= 1 + A_T^{(\phi_T - 1)} f_T(x) + A_E^{(\phi_E - 1)} f_E(x) + A_T^{(\phi_T + \phi_E - 2)} f_T(x) f_E(x)
\]

which presents a modestly complicated dynamic structure. This proves the proposition.

**Proposition 4** The “naïve” regression system (P3.4) is not stable but can be saddle-path stable.

**Proof:**

Given that

\[
\begin{align*}
(M_T - 1) & \approx \frac{\dot{A}_T}{A_T} = A_T^{(\phi_T - 1)} f_T(x) \\
(M_E - 1) & \approx \frac{\dot{A}_E}{A_E} = A_T^{(\phi_E - 1)} f_E(x)
\end{align*}
\]
Define $A = A_r \cdot A_E$, then
\[
\begin{align*}
\dot{A}_r &= A_r^T A_E^{(r-1)} f_r(x) \\
\dot{A}_E &= A_r^{(r-1)} A_E^T f_E(x)
\end{align*}
\] (P4.1)

To simplify the derivation we assume $f(1) = 1$. Taylor linearization at the point $1$ leads to
\[
\begin{align*}
\dot{A}_r &\approx \delta_{T0} + \phi A_r + (\phi-1) A_r + \sum \beta_r x_i \\
\dot{A}_E &\approx \delta_{E0} + (\theta - 1) A_r + \theta A_r + \sum \beta_E x_i
\end{align*}
\] (P4.2)

(P4.2) forms a system of ordinary differential equations. For stability of the system, it requires that the two eigenvalues of the parameter matrix
\[
\begin{bmatrix}
(\phi - r) & (\phi - 1) \\
(\theta - 1) & (\theta - r)
\end{bmatrix}
\] (P4.3)
to be real and negative. The determinant of the characteristic matrix is given by
\[
\begin{vmatrix}
(\phi - r) & (\phi - 1) \\
(\theta - 1) & (\theta - r)
\end{vmatrix} = 0
\] (P4.4)

Rewrite slightly
\[
\begin{vmatrix}
(1 - r) + (\phi - 1) & (\phi - 1) \\
(\theta - 1) & (1 - r) + (\theta - 1)
\end{vmatrix} = 0
\] (P4.5)

Let $R = (1 - r)$, $\phi' = \phi - 1$, and $\theta' = \theta - 1$, then
\[
\begin{vmatrix}
R + \phi' & \phi' \\
\theta' & R + \theta'
\end{vmatrix} = 0
\] (P4.6)

\[
(R + \phi')(R + \theta') - \phi' \theta'
= [R^2 + (\phi' + \theta')R + \phi' \theta'] - \phi' \theta'
= R^2 + (\phi' + \theta')R = 0
\]

$R_1 = 0, r_1 = 1$

$R_2 = -(\phi' + \theta')$

$r_2 = \phi' + \theta' + 1 = \phi - 1 + \theta - 1 + 1' = \phi + \theta - 1$

16
Therefore, for saddle path stable solution, it requires \((\phi + \theta) < 1\) or \(\phi < (1- \theta)\). This proposition allows *Efficiency change equation and technical progress equation to have different structures with respect to the level of technology.* But it does not rule out the possibility of a similar structure for the two equations. For example, if we apply the structure of the knowledge production function (3.5) to technical progress index (4.6), with \(\phi < 0\) the higher the level of technology the slower the speed of technical progress. The structure implied in the formulation of Benhabib and Spiegel (1994) is somewhat similar and pointing to a similar direction. The lower the level of technical efficiency, the stronger the incentive for catching up with the best practice, i.e., \(\theta < 0\). An empirical example of this type of results can be found in Zheng (2008). We next take a look at the more general case of (P3.5).

**Proposition 5** The system (P3.5) and (P3.3) can be stable.

Proof:

Since we have

\[
\begin{align*}
\dot{A}_T &= A_T^\phi \ A_E^\theta \ f_T(x) \\
\dot{A}_E &= A_T^\phi \ A_E^\theta \ f_E(x)
\end{align*}
\]

(P3.5)

Taylor linearization leads to

\[
\begin{align*}
\dot{A}_T &\approx \delta_T + \phi_T A_T + \theta_T A_E + \sum \beta_T x_i \\
\dot{A}_E &\approx \delta_E + \phi_E A_T + \theta_E A_E + \sum \beta_E x_i
\end{align*}
\]

(P5.6)

For stability of the system, it requires that the two eigenvalues of the parameter matrix

\[
\begin{bmatrix}
\phi_T & \theta_T \\
\phi_E & \theta_E
\end{bmatrix}
\]

(P5.7)

to be real and negative. The determinant of the characteristic matrix is given by

\[
\begin{vmatrix}
\phi_T - r & \theta_T \\
\phi_E & \theta_E - r
\end{vmatrix} = 0
\]

(P5.8)

where \(r\) is the characteristic root. This means:
\((\phi_f - r)(\theta_E - r) - \theta_f \phi_E = 0\) \hfill (P5.9)

\[ r^2 - (\phi_f + \theta_E) r + (\phi_f \theta_E - \theta_f \phi_E) = 0 \]

\[ r = \frac{-(\phi_f + \theta_E) \pm \sqrt{(\phi_f + \theta_E)^2 - 4(\phi_f \theta_E - \theta_f \phi_E)}}{2} \] \hfill (P5.10)

After Simplification

\[ r = \frac{\phi_f + \theta_E \pm \sqrt{(\phi_f - \theta_E)^2 + 4\theta_f \phi_E}}{2} \] \hfill (P5.11)

For stability it is required that

\((\phi_f - \theta_E)^2 + 4\theta_f \phi_E \geq 0\) \hfill (P5.12)

and

\[ r = \frac{\phi_f + \theta_E \pm \sqrt{(\phi_f - \theta_E)^2 + 4\theta_f \phi_E}}{2} < 0 \] \hfill (P5.13)

which ends our proof for Proposition 5.

If we apply Proposition 5 to the simplest scenario described in (P3.3), with \(\phi_E = \theta_f = 0\) (P5.11) reduces to

\[ r = \frac{\phi_f + \theta_E \pm (\phi_f - \theta_E)}{2} \]

so

\[ r_1 = \phi_f = \phi \text{ and } r_2 = \theta_f = \theta \]

For system stability it requires that both eigenvalues to be negative and saddle-path stability is also possible in this simple case.

To further simplify the system, one alternative might be to make the decomposition of Malmquist index additive instead of multiplicative by taking logarithm of both sides of (4.2).
However, since log cannot be used for negative values of Malmquist TFP index, we pursue a linearization of (4.2) using the first-order Taylor expansion at the point zero as follows.\(^6\)

**Proposition 6** (P3.2) can be conveniently linearized.

Proof:
We have already seen this in (4.16) from a slightly different perspective. Expand (P3.2) around the point of (0,0) with respect to growth rates in \(A_T\) and \(A_E\) yields,

\[
M_0 = M_T \cdot M_E \approx \frac{A_T + \dot{A}_T}{A_T} + \frac{A_E + \dot{A}_E}{A_E} = \left( \frac{\dot{A}_T}{A_T} + 1 \right) \left( \frac{\dot{A}_E}{A_E} + 1 \right)
\]

\(\text{(P6.1)}\)

For example, if we refer to specification in (P3.4), we have

\[
M_0 - 1 = \frac{\dot{A}}{A} \approx A^{(\phi-1)} f_T(x) + A^{(\theta-1)} f_E(x)
\]

\(\text{(P6.2)}\)

and

\[
\dot{A} \approx A^\phi f_T(x) + A^\theta f_E(x)
\]

\(\text{(P6.3)}\)

which is rather convenient and consistent with the Nelson and Phelps (1966) theoretical and Benhabib and Spiegel (1994) empirical formulations.

We propose to treat the additive version of the productivity decomposition in Benhabib and Spiegel (1992), which was an empirical reformulation of Nelson and Phelps (1966). However, since the Malmquist index decomposition is multiplicative, we here suggest to linearize (P3.5) directly as follows:

\[
\begin{align*}
\frac{\dot{A}_T}{A_T} &\approx \delta_{T_0} + \delta_T (\phi_T - 1) A_T + \delta_T \theta_T A_E + \delta_T \sum \beta_{Ti} x_i \\
\frac{\dot{A}_E}{A_E} &\approx \delta_{E_0} + \delta_E \phi_E A_T + \delta_E (\theta_E - 1) A_E + \delta_E \sum \beta_{Ei} x_i
\end{align*}
\]

\(\text{(P6.4)}\)

For simplicity we set \(\delta = 1\). According to (P6.1)

---

\(^6\) According to the basic assumption A5 TFP growth cannot be negative, so log linearization in principle should be allowed. But in practice, it’s better to be able to accommodate negative values of TFP change or growth at this stage. Taking positive values restricts data to those observations that can only be generated with a well-defined knowledge production function.
Then we have
\[
\frac{\dot{A}}{A} \approx \frac{\dot{A}_T}{A_T} + \frac{\dot{A}_E}{A_E} + \left( \frac{\delta_{\phi}(\phi_E - 1)}{A_T} + \frac{\delta_{\theta}(\theta_E - 1)}{A_E} \right) x_i
\]
\[
\text{(P6.5)}
\]

Taking account (P3.5), we have
\[
\begin{align*}
\frac{\dot{A}}{A} & \approx \left( \delta_{T0} + \delta_{E0} \right) + \left[ \delta_{\phi} + \delta_{\phi} (\phi_E - 1) \right] A_T + \left[ \delta_{\theta} + \delta_{\theta} (\theta_E - 1) \right] A_E \\
& \quad + \sum \left( \delta_{\phi} \beta_i + \delta_{\phi} \beta_i \right) x_i \\
\frac{\dot{A}_T}{A_T} & \approx \delta_{T0} + \delta_{T0} A_T + \delta \sum \beta_i x_i \\
\frac{\dot{A}_E}{A_E} & \approx \delta_{E0} + \delta_{E0} A_E + \delta \sum \beta_i x_i
\end{align*}
\]
\[
\text{(P6.6)}
\]

which is a linear dynamic system with linear cross-equation constraints. The advantage of cross-equation linear constraints is that one can impose these constraints in linear regressions to achieve better efficient parameter estimations. For example, in Zheng, Liu and Bigsten (2003), TFP growth and its components were used in regressions on the same set of the explanatory variables. Although no cross-equation restrictions were applied there, one can see in their Table 5 that the sum of the parameters in the two TFP growth component equations was very close to the parameter estimate in the TFP growth equation. Not all parameter estimates were statistically significant without imposing cross-equation restrictions. The situation may change if one does so.

5. An Empirical Example of Norwegian Firms
We now update the results in Zheng (2008) in which a Norwegian firm panel and CIS3 and CIS4 data are used jointly. The complete set of the six regression equations is listed as follows:

(1) Technical efficiency
(2) Best practice (binary)
We thus use technical efficiency, technical efficiency change, technical progress as dependent variables; and level TFP, innovation intensity, and measure of R&D personnel are used as independent variables together with other auxiliary variables of the CIS datasets. We will focus on the two Malmquist index component equations as follows:

\[
\text{EFFCH} = \alpha_2 + \delta_2 (\theta - 1)\text{TFP} + \delta_2 \lambda_2 L_A + \rho_2 \frac{\Delta R}{Y} \\
\]

\[
\alpha_2 = (1 + \delta_2) - \delta_2 (\theta - 1) - \delta_2 \lambda_2 \\
\]

\[
\text{TECH} = \alpha_3 + \delta_3 (\phi - 1)\text{TFP} + \delta_3 \lambda_3 L_A + \rho_3 \frac{\Delta R}{Y} \\
\]

\[
\alpha_3 = (1 + \delta_3) - \delta_3 (\phi - 1) - \delta_3 \lambda_3 \\
\]

As in Zheng (2008), (6.1) and (6.2), with all parameters identified, are the basic microstructure we apply for the regression analyses. Note that since \( L_A = sL \) both the ratio of researchers to the total employed and the total labor can be included in a regression. This, on one hand, can make the magnitude between explanatory variables more comparable, and on the other hand can be used to test if the parameter associated with the ratio is statistically different from that associated from the total labor \( L \). In empirical estimations, in the case of labor input sometimes it is better to use a unit of measure that is comparable with the productivity index on the right hand side of the equation. For example labor can be standardized using a “mean transformation” or log transformation. In Zheng (2008) log transformation was used, but in the updated results in Table 1 we have used the “mean transformation.”
**Table 1** Regression results for the Malmquist dynamic regression  
(Norwegian Firms 1995-2004)

<table>
<thead>
<tr>
<th></th>
<th>Best practice</th>
<th>Efficiency Change</th>
<th>Technical progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-3.39***</td>
<td>0.419***</td>
<td>1.829***</td>
</tr>
<tr>
<td></td>
<td>(-5.84)</td>
<td>(2.73)</td>
<td>(21.26)</td>
</tr>
<tr>
<td>TFP level (A)</td>
<td></td>
<td>0.631***</td>
<td>-0.547***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(6.11)</td>
<td>(-8.50)</td>
</tr>
<tr>
<td>Employment</td>
<td>0.072***</td>
<td>-0.008</td>
<td>-0.009*</td>
</tr>
<tr>
<td></td>
<td>(3.79)</td>
<td>(-0.90)</td>
<td>(-1.81)</td>
</tr>
<tr>
<td>R&amp;D Personnel</td>
<td>2.02***</td>
<td>0.045</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>(2.83)</td>
<td>(0.22)</td>
<td>(0.20)</td>
</tr>
<tr>
<td>Innovation intensity</td>
<td>-0.749</td>
<td>0.045</td>
<td>-0.196</td>
</tr>
<tr>
<td></td>
<td>(-0.84)</td>
<td>(0.21)</td>
<td>(-1.35)</td>
</tr>
<tr>
<td>R&amp;D Purchase intensity</td>
<td>0.376</td>
<td>0.200*</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.89)</td>
<td>(1.68)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Patent</td>
<td>0.222</td>
<td>-0.006</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>(1.16)</td>
<td>(-0.10)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>Export intensity</td>
<td>0.338*</td>
<td>-0.029</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>(1.69)</td>
<td>(-0.43)</td>
<td>(0.58)</td>
</tr>
<tr>
<td>Continuous R&amp;D</td>
<td>-0.213</td>
<td>-0.073**</td>
<td>-0.026</td>
</tr>
<tr>
<td></td>
<td>(-1.47)</td>
<td>(-2.10)</td>
<td>(-1.28)</td>
</tr>
<tr>
<td>Market orientation</td>
<td>0.178*</td>
<td>0.018</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(1.90)</td>
<td>(0.68)</td>
<td>(-0.28)</td>
</tr>
<tr>
<td>Product life</td>
<td>-0.011</td>
<td>-0.011</td>
<td>-0.018</td>
</tr>
<tr>
<td></td>
<td>(-0.14)</td>
<td>(-0.58)</td>
<td>(-1.64)</td>
</tr>
<tr>
<td>Market location</td>
<td>0.091</td>
<td>-0.026</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(1.06)</td>
<td>(-0.97)</td>
<td>(0.14)</td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>96.06</td>
<td>(Sigma) 0.376</td>
<td>(Sigma) 0.376</td>
</tr>
<tr>
<td>P-value</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>939</td>
<td>923</td>
<td>923</td>
</tr>
</tbody>
</table>
The result in Table 1 gives the impression that there exists increasing returns in the production of the new ideas since the implied $\theta$ might be in the range of 2 to 3 according to the calculation based on (6.1). However, since this increasing returns effect is only present for the technical efficiency component of the TFP growth, then in the long run without any effect of knowledge stock on technical progress, the increasing returns to the production of new ideas will cease to exist. It is therefore important to distinguish the part of TFP growth due to technical efficiency improvement from the part that resulted from technical progress in both theoretical and empirical analysis.

Although the sign for the parameter estimate associated with technology level in Table 1 appears to be the right one, the implied $\phi$ could be as large as -10, indicating considerable decreasing returns to the production of new ideas. Combining our estimate of $\phi$ and $\theta$, the dynamic system is saddle path stable according to Proposition 3. Whether or not this is a reasonable magnitude would require more applications of this kind to emerge in the future. We can also try simulations to check the range of the parameters.

6. Summary

The DEA-based Malmquist productivity index model has become increasingly popular in empirical estimations since its introduction to the literature about two decades ago. However, regression analyses based on the Malmquist productivity index and its components are far less in numbers than those based on straight estimates of DEA efficiency scores. As the progresses are being made with regard to the understanding of the statistical properties of the two-stage procedure, the need for a systematic treatment of the Malmquist index regression equations has become an urgent task. In this study we have approached the issue through the knowledge production function framework suggested and constructed the multiple regression equations into an interactive, interrelated, and fully specified dynamic system. The immediate new knowledge acquired through this process is that the system consisting of the “naïve” Malmquist regression equations can be saddle-path stable.

On the other hand, the dynamic system constructed in this study in general has three attractive properties. First, the Malmquist index is consistent with the Solow residual under rather standard assumptions, which paves the way for developing the set of Malmquist index regression equations into a dynamic system under the framework of the knowledge production function. Second, cross-equation restrictions not only exist but also can be imposed in empirical estimations using a linearised version of the Malmquist index. Third, the dynamic nature of the equation system widens the scope for the knowledge production function to be specified. As we have mentioned in the study, the endogenous growth literature basically says nothing about the productivity decomposition with respect to technical progress on the one hand and technical efficiency improvement on the other. But our specifications of the knowledge production function reveal that several different systems can be considered;
and knowledge stocks can be further classified into those that may be effective in improving technical efficiency and those that might help push forward the production frontier, i.e., the utilization of new technologies. This property of the system enriches the structures that can be informative of the knowledge acquisition process, and should have important implications for micro studies on the relationships between innovation and productivity and for macro modeling of endogenous economic growth.
References


