# Aggregation of Malmquist Productivity Indexes

Zelenyuk, Valentin

Institut De Statistique, UniversitE Catholique De Louvain

May 2005

**INSTITUT DE STATISTIQUE**

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

**DISCUSSION PAPER**

**0505**

# AGGREGATION OF MALMQUIST PRODUCTIVITY INDEXES

ZELENYUK, V.

http://www.stat.ucl.ac.be

# Aggregation of Malmquist Productivity Indexes

**Valentin Zelenyuk***

**Abstract**

In this paper we extend the work of Färe and Zelenyuk (2003) to find a theoretically justified method of aggregating Malmquist Productivity Indexes over individual decision making units (firms, countries, etc.) into a group Malmquist Productivity Index. We also consider the aggregation of decomposed parts of the Malmquist Productivity Index to obtain a decomposition of the Malmquist Productivity Index for a particular group.

**Key Words:** DEA, Efficiency, Productivity, Index aggregation.

**Feb 8, 2004**

* Institut de Statistique, Université Catholique de Louvain, Belgium and UPEG (Ukrainian Productivity and Efficiency Group) at EERC of "Kyiv-Mohyla Academy", Kyiv (Kiev), Ukraine. *Mailing address:* 20 voie du Roman Pays, B 1348, Louvain-la-Neuve, Belgium. Tel. : +32-10-474311, fax : +32-10-473032. zelenyuk@stat.ucl.ac.be

**Acknowledgement:** I would like to thank Tom Coupé, Rolf Färe, Kevin Fox, Roy Gardner, Daniel Henderson, Natalia Konstandina, Vladimir Nesterenko, and Mykhaylo Salnykov for their valuable comments. I also would like to thank participants of NAPW 2004 (University of Toronto, Toronto, Canada) for their valuable comments. I would like to give special thanks to two anonymous referees who helped improve the clarity of this paper. Research support from the "Interuniversity Attraction Pole", Phase V (No.P5/24) from the Belgian Government (Belgian Science Policy) is acknowledged. Any views and remaining errors are my responsibility.

## 1. Introduction

One of the most popular approaches to measuring productivity changes is based on using Malmquist Productivity Indexes—a method originated by Caves et al. (1982). Virtually any empirical study that uses this approach, at some point, reports averages of the productivity indexes they estimate—to represent the overall tendency in productivity changes, make inference, etc. Most of the time, the *equally-weighted geometric* mean is used for this purpose. Recent developments on aggregation in a closely related field—efficiency analysis—have emphasized the importance of using *weights* in the aggregation of *indexes*. The purpose of this is to account for the relative importance of each observation whose efficiency score is entering into the average (Färe and Zelenyuk, 2003). It is quite natural to look at this same issue in the context of productivity indexes, as we do in this paper.

The rest of the paper is structured as follows. Section 2 outlines the key definitions for the individual (disaggregated) case, while Section 3 does this for the group (aggregated) case. The main aggregation result is given in Section 4, and the resulting aggregate productivity index is then compared, in Section 5, to the commonly used geometric average. Section 6 demonstrates how the same aggregation principle can be applied to the context of decompositions of productivity indexes into different sources of change (e.g., efficiency and technology change). Finally, Section 7 lists some possible extensions for further research as well as concludes this study.

## 2. Characterization of Technologies and Measurement of Productivity Changes

For each DMU (decision making unit: plant, firm, country, etc.) $k$ ($k = 1, 2,\ldots, n$), let $x_\tau^k = (x_{\tau,1}^k,\ldots,x_{\tau,N}^k)' \in \Re_+^N$ be a vector of $N$ *inputs* that DMU $k$ uses in period $\tau$ (for our case, $\tau = s, t$) to produce a vector of $M$ *outputs*, denoted by $y_\tau^k = (y_{\tau,1}^k,\ldots,y_{\tau,M}^k)' \in \Re_+^M$. We assume the technology of DMU $k$ in a period $\tau$ is characterized by the output set

$$P_\tau^k(x^k) \equiv \{y \; : \; y \in \Re_+^M \; is \; producible \; from \; x^k \in \Re_+^N\}. \tag{2.1}$$

Throughout, we assume the technology in any period $\tau$ satisfies the usual regularity axioms of production theory.[1] Thus we can use the *output oriented* Shephard's (1970) distance function $D_\tau^k : \Re_+^N \times \Re_+^M \to \Re_+^1 \cup \{\infty\}$, defined as

$$D_\tau^k(x^k, y^k) \equiv \inf\{\theta : \ y^k / \theta \in P_\tau^k(x^k)\}, \tag{2.2}$$

to obtain a *complete* (*primal*) *characterization of the technology* of DMU *k* in period $\tau$, in the sense that

$$D_\tau^k(x^k, y^k) \le 1 \quad \Leftrightarrow \quad y^k \in P_\tau^k(x^k). \tag{2.3}$$

This function is also a convenient criterion for measuring the relative distance from any input-output combination of some DMU *k* towards the frontier of the technology set. In particular, if we let the technology frontier of $P_\tau^k(x^k)$, $x^k \in \Re_+^N$ be defined (for period $\tau$) as

$$\partial P_\tau^k(x^k) = \{y \in \Re_+^M \quad : \quad y \in P_\tau^k(x^k), \quad \lambda y \notin P_\tau^k(x^k), \quad \forall \lambda \in (1, +\infty)\},$$

then,

$$0 < D_\tau^k(x^k, y^k) < 1 \quad \Leftrightarrow \quad y^k \in P_\tau^k(x^k),\ y^k \notin \partial P_\tau^k(x^k),\ y^k \ne 0,$$

$$D_\tau^k(x^k, y^k) > 1 \quad \Leftrightarrow \quad y^k \notin P_\tau^k(x^k), \quad \exists \lambda > 0 : \ \lambda y^k \in P_\tau^k(x^k)\ ,$$

$$D_\tau^k(x^k, y^k) = 1 \quad \Leftrightarrow \quad y^k \in \partial P_\tau^k(x^k)\ ,$$

and

$$D_\tau^k(x^k, y^k) = 0 \quad \Leftrightarrow \quad y^k = 0.$$

These (and other[2]) properties have made the function (2.2) very popular in efficiency analysis, where it can be used to define the Farrell (1957) *technical efficiency* measure of DMU *k* (in period $\tau$) as

$$TE_\tau^k(x^k, y^k) = 1 / D_\tau^k(x^k, y^k),$$

An alternative (*dual*) characterization of $P_s^k(x^k)$ can be given via the *revenue* function,

---

[1] We assume all the output sets satisfy *free disposability of outputs*, i.e., $y^o \in P_\tau^k(x) \Rightarrow y \in P_\tau^k(x), \forall y \leqq y^o$ and are compact (for all $x \in \Re_+^N$). We also assume $y \notin P_\tau^k(\mathbf{0}_N)$, $\forall y \ge \mathbf{0}_M$ ("*no free lunch*") and $\mathbf{0}_M \in P_\tau^k(x)$, $x \in \Re_+^N$ ("*producing nothing is possible*"). See Färe and Primont (1995) for details.

[2] For the list of properties of the distance function and proofs, see Shephard (1970) and Russell (1990, 1997).

$$R_\tau^k(x^k, p) \equiv \max_y\{p y : y \in P_\tau^k(x^k)\}, \tag{2.4}$$

where $p = (p_1, ..., p_M) \in \Re_+^M$ denotes the vector of output prices.[3] Assuming $py^k \neq 0$, the revenue function can be used to define the measure of *revenue efficiency* of DMU *k* in period $\tau$, as

$$RE_\tau^k(x^k, y^k, p) \equiv R_\tau^k(x^k, p) / py^k. \tag{2.5}$$

The revenue function is *dual* to the distance function $D_\tau^k(x^k, y^k)$, in the sense that [4]

$$D_\tau^k(x^k, y^k) = \sup_p\{py^k / R_\tau^k(x^k, p)\}. \tag{2.6}$$

This expression implies that the revenue efficiency measure is an upper bound to the technical efficiency measure, i.e.,

$$R_\tau^k(x^k, p) / p y^k \geq 1 / D_\tau^k(x^k, y^k). \tag{2.7}$$

This statement, also known as Mahler's inequality, is often used to define, in a residual fashion, the measure of *allocative efficiency* of DMU *k* (for period τ)

$$AE_\tau^k(x^k, y^k, p) \equiv RE_\tau^k(x^k, y^k, p) \times D_\tau^k(x^k, y^k), \tag{2.8}$$

The idea of decomposition (2.8) goes back at least to Farrell (1957) and will prove very useful in deriving our aggregation results.

Let us now turn to the measurement of productivity changes from one period (*s*) to another period (*t*). The Shephard's distance function, defined above, is often used for defining Malmquist productivity indexes. This concept was first introduced by Caves et al. (1982), who suggested two indexes that differ with respect to the reference technology they are measured to. In general, such two

[3] For the purpose of obtaining the desired aggregation results we have made a *necessary* assumption that all firms face the same *output* prices.
[4] To achieve this result, convexity of the output sets is needed, in addition to other regularity axioms mentioned above; see Färe and Primont (1995) for details.

indexes are not equal, and a common practice is to avoid arbitrariness of choice by taking the geometric mean to define the (output oriented) Malmquist Productivity Index (further MPI) as

$$M^k(y_s^k, y_t^k, x_s^k, x_t^k) \equiv \left[\frac{D_s^k(x_t^k, y_t^k)}{D_s^k(x_s^k, y_s^k)} \times \frac{D_t^k(x_t^k, y_t^k)}{D_t^k(x_s^k, y_s^k)}\right]^{1/2}. \tag{2.9}$$

Note that now we have time subscripts for the output and input vectors. This indicates that we now look at specific values of the Shephard's distance functions, evaluated at the actual input-output allocations, $(x_\tau^k, y_\tau^k)$, for particular DMU $k$ ($k=1,\ldots,n$), in the particular period $\tau$ ($\tau = s, t$).

If we take the revenue efficiency measure (2.5) and recall its dual relationship to the distance function via (2.6) or (2.7), we can define the *revenue* (or dual) analog of the *MPI* as

$$RM^k(\cdot) \equiv RM^k(p_s, p_t, y_s^k, y_t^k, x_s^k, x_t^k) \equiv \left[\left(\frac{RE_s^k(x_t^k, y_t^k, p_t)}{RE_s^k(x_s^k, y_s^k, p_s)} \times \frac{RE_t^k(x_t^k, y_t^k, p_t)}{RE_t^k(x_s^k, y_s^k, p_s)}\right)^{-1}\right]^{1/2}, \tag{2.10}$$

which can also be decomposed (analogously to (2.8), but for the context of productivity indexes) as

$$RM^k(\cdot) \equiv M^k(\cdot) \times AM^k(\cdot), \tag{2.11}$$

where $M^k(\cdot)$ is given in (2.9), and

$$AM^k(\cdot) \equiv AM^k(p_s, p_t, y_s^k, y_t^k, x_s^k, x_t^k) \equiv \left[\left(\frac{AE_s^k(x_t^k, y_t^k, p_t)}{AE_s^k(x_s^k, y_s^k, p_s)} \times \frac{AE_t^k(x_t^k, y_t^k, p_t)}{AE_t^k(x_s^k, y_s^k, p_s)}\right)^{-1}\right]^{1/2}. \tag{2.12}$$

In the next section we define the *group* analogs of these *individual* indexes.

## 3. *Group* Efficiency and Productivity Measures

We will denote the input and output allocation among DMUs *within* the group, respectively, by $X_\tau = (x_\tau^1,\ldots,x_\tau^n)$ and $Y_\tau = (y_\tau^1,\ldots,y_\tau^n)$, we will also denote the *sum* of output vectors over all DMUs in the group with $\overline{Y}_\tau = \sum_{k=1}^n y_\tau^k$ ($\tau = s, t$).

A critical step is to define a *group technology*—the *aggregate* technology of *all* DMUs within the group. In the context we have chosen—*output orientation* (i.e., consideration of output changes *given fixed* levels of inputs)—a natural way to define the group technology is to assume the *additive structure of aggregation* of the *output sets* (Färe and Zelenyuk, 2003), i.e.,

$$\overline{P}_{\tau}(X) \equiv \sum_{k=1}^{n} P_{\tau}^{k}(x^{k}), \qquad \tau = s, t. \tag{3.1}$$

Thus, the *output set* of a group of DMUs, $\overline{P}_{\tau}(X)$, is the sum of the individual *output sets* of all DMUs in this group. The properties of this group technology depend on the properties of technologies of each DMU in the group. In particular, $\overline{P}_{\tau}(X)$ inherits the regularity conditions we assumed above and is convex if the individual output sets are convex.

Given the group technology (3.1), the *group revenue* function can be defined as

$$\overline{R}_{\tau}(X, \; p) \equiv \max_{y}\{py \;\; : \;\; y \in \overline{P}_{\tau}(X)\}, \qquad \tau = s, t. \tag{3.2}$$

which is a group analog to (2.4) and the *group* analog of (2.5) can be defined as

$$\overline{RE}_{\tau}(X, \overline{Y}, p) \equiv \overline{R}_{\tau}(X, p) / p\overline{Y}, \qquad \tau = s, t. \tag{3.3}$$

In the context of measuring productivity changes between period *s* and *t*, we can now define the *group* or aggregate analog of (2.10) as

$$\overline{RM}(p_{s}, p_{t}, \overline{Y}_{s}, \overline{Y}_{t}, X_{s}, X_{t}) \equiv \left[ \left( \frac{\overline{RE}_{s}(X_{t}, \overline{Y}_{t}, p_{t})}{\overline{RE}_{s}(X_{s}, \overline{Y}_{s}, p_{s})} \times \frac{\overline{RE}_{t}(X_{t}, \overline{Y}_{t}, p_{t})}{\overline{RE}_{t}(X_{s}, \overline{Y}_{s}, p_{s})} \right)^{-1} \right]^{1/2}, \tag{3.4}$$

where, again, the time subscripts indicate that we now look at the specific values of efficiency measures realized in particular periods $\tau$ ($\tau = s, t$).

Our goal is to find an aggregation function (especially the aggregation weights) $f_{R}(\cdot)$ that would relate (3.4) to the individual measures (2.10) or its components given by (2.5), for all *k*, i.e., establish

$$\overline{RM}(p_s, p_t, \overline{Y_s}, \overline{Y_t}, X_s, X_t) = f_R(RE_\tau^1(\cdot),..., RE_\tau^n(\cdot)), \qquad \tau = s, t. \tag{3.5}$$

so that, *preferably*, we maintain the decomposition (2.11) at the *aggregate* level. Formally, we want

$$\overline{RM}(p_s, p_t, \overline{Y_s}, \overline{Y_t}, X_s, X_t) = \overline{M}(\cdot) \times \overline{AM}(\cdot). \tag{3.6}$$

where, again we need to find some aggregation functions $f_T(\cdot), f_A(\cdot)$, so that the aggregate *primal* MPI can be obtained from the individual analogs (2.9) or its components given by (2.2), i.e.,

$$\overline{M}(\cdot) \equiv \overline{M}(\overline{Y}_s, \overline{Y}_t, X_s, X_t) \equiv f_T(D_\tau^1(\cdot),..., D_\tau^n(\cdot)), \qquad \tau = s, t. \tag{3.7}$$

Further, the aggregate *allocative* MPI can be obtained from (2.12) or its components, given by (2.8), i.e.,

$$\overline{AM}(\cdot) \equiv \overline{AM}(\overline{Y}_s, \overline{Y}_t, X_s, X_t) \equiv f_A(AE_\tau^1(\cdot),..., AE_\tau^n(\cdot)), \qquad \tau = s, t. \tag{3.8}$$

In the next section we will find such functions.

## 4. Aggregation Results for the Malmquist Productivity Indexes

The fundamental result for our study is an intertemporal extension of the result derived by Färe and Zelenyuk (2003). Specifically, it says

$$\overline{R}_\tau(X, p) = \sum_{k=1}^n R_\tau^k(x^k, p), \qquad x^k \in \Re_+^N, \forall k = 1,...,n, \ p \in \Re_+^M, \quad \tau = s, t. \tag{4.1}$$

The economic intuition of this theorem is straightforward. The sum of the revenues of *individual* revenue-maximizing DMUs in a given group is the same as the revenue obtained by a revenue-maximizing union of these DMUs (e.g., an industry or its sub-groups, regions such as APEC, EU, NAFTA, etc.) whose technology is defined in (3.1) and assuming that the output price vector is the same for all DMUs. [5] (The proof of this result is essentially the same as the proof of Färe and Zelenyuk (2003) and therefore is skipped for the sake of brevity). Considering the context of measuring productivity changes between periods *s* and *t*, using (4.1), we obtain the key expression for our study

$$\overline{RE_{\tau}}(X_j, \overline{Y}_j, p_j) = \sum_{k=1}^{n} RE_{\tau}^{k}(x_j^k, y_j^k, p_j) \cdot S_j^k, \qquad j, \tau = s, t. \tag{4.2}$$

where

$$S_j^k \equiv p_j y_j^k / p_j \overline{Y}_j, \qquad k = 1, \ldots, n; \qquad j = s, t. \tag{4.3}$$

Moreover, after a little more of algebra, we obtain the desired decomposition,

$$\overline{RE_{\tau}}(X_j, \overline{Y}_j, p_j) = \overline{TE_{\tau}}(j) \times \overline{AE_{\tau}}(j), \qquad j, \tau = s, t \tag{4.4}$$

where

$$\overline{TE_{\tau}}(j) \equiv \sum_{k=1}^{n} [D_{\tau}^{k}(x_j^k, y_j^k)]^{-1} \cdot S_j^k, \qquad j, \tau = s, t. \tag{4.5}$$

$$\overline{AE_{\tau}}(j) \equiv \sum_{k=1}^{n} AE_{\tau}^{k}(x_j^k, y_j^k, p_j) \cdot S_{ae,j}^k, \qquad j, \tau = s, t. \tag{4.6}$$

and

$$S_{ae,j}^k \equiv \frac{p_j \left(y_j^k / D_{\tau}^k(x_j^k, y_j^k)\right)}{p_j \sum_{k=1}^{n} \left(y_j^k / D_{\tau}^k(x_j^k, y_j^k)\right)}, \qquad k = 1, \ldots, n; \quad j, \tau = s, t. \tag{4.7}$$

*Remark 1*. The proposed weights of aggregation for obtaining the group measures are *not* ad hoc but are derived from economic principles (agents' optimization behavior). Incidentally, the weights are also quite intuitive and perhaps resemble what common sense would suggest. They account for the importance of the technical and revenue efficiency scores of DMUs via the share of the value of total output of these DMUs in the group. The weights for aggregating allocative efficiency are similar, but the technically efficient output is used instead of the actual one. This is also what one might logically expect. It is also worth noting that in a single output case, the weights clearly reduce to the output shares—exactly the weights proposed by Farrell (1957) while envisioning his concept of the *Structural* Efficiency of an Industry. Such context of a single output is not uncommon in practice. In cross-country efficiency and productivity analysis, for example, researchers often proxy all outputs with one aggregate output, e.g., GDP, in which case the weights would be the GDP-shares (e.g., see Henderson and Zelenyuk, 2004).

---

[5] This theorem is a revenue analog to the Koopmans (1957) theorem of aggregation of the profit functions. The cost analog is proven in Färe, Grosskopf and Zelenyuk (2002a).

*Remark 2*. Interestingly, a similar weighting scheme has been suggested by Domar (1961)—for the context of aggregation across industries, however they were obtained under different (more restrictive) assumptions and using a different derivation method than ours.

*Remark 3*. The weights of aggregation *depend* on prices. This shall not be surprising given the fact that for the derivation of these weights (and the aggregation function) an *economic* criterion was used—revenue optimization.—and, loosely speaking, 'the quintessence of economics is in prices.' Importantly, the prices are required to be the same for all DMUs (firms, countries, etc). On the one hand, this type of "Law of One Price" assumption is seldom true in reality (at least some statistical noise might be present) and in fact has received a considerable attention and critique in the recent literature (e.g., see Cherchye et al. (2004) and references therein). On the other hand, such an assumption is consistent with many standard economic models of perfect competition, Cournot-type oligopoly, monopolistic competition, etc. In any case, to our knowledge, the scheme (4.2)-(4.4) is so far the only *positive* aggregation result and the "Law of One Price" assumption is a *necessary* condition for it to hold, but we hope future research would relax this assumption. Meanwhile, for an empirical researcher, this assumption might be a simplifying one that gives a way of summarizing a large number of obtained efficiency (and, as later will be shown, productivity) scores into one number representing an entire economic system (or its sample), so that the economic importance of each unit in this system is accounted for with a theoretically justified and an intuitive weight. For within industry studies one could use, for example, the average prices, which are often publicly available. For cross-country analysis, for example, the world prices might be used.

*Remark 4*. If price information is unavailable then an empirical researcher might have to call for some additional estimation or/and simplifications. For example, one could resort to (estimated) shadow prices for the entire system (or take an average of individual shadow prices). Alternatively, researchers might be willing to accept an additional standardization, e.g., of the type proposed by Färe and Zelenyuk (2003), for making the weights derived above price-independent, while still preserving the aggregation structure based on economic optimization criterion. Specifically, one may assume that

$$\frac{p_{j,m}\overline{Y}_{j,m}}{\sum_{m=1}^{M} p_{j,m}\overline{Y}_{j,m}} = a_{j,m}, \qquad m = 1, \ldots, M; \qquad j = s, t, \tag{4.8}$$

where $\overline{Y}_{j,m} \equiv \sum_{k=1}^{n} y_{j,m}^{k}$, and $a_{j,m}$ ($m = 1, \ldots, M$) are assumed to be known (or estimated) constants between zero and unity, so that $\sum_{m=1}^{M} a_{j,m} = 1$. In words, expression (4.8) says that (in period $j$) the *share* of the industry revenue from output $m$ in the industry *total* revenue is equal to some constant $a_{j,m}$. In practice, such aggregate information on the value shares of each output in an industry, $a_{j,m}$, is often available from industry surveys, governmental reports or previous studies.[6] Imposing (4.8) onto the weights for aggregating technical (and revenue) efficiencies yields price independent weights,

$$\tilde{S}_j^k = \sum_{m=1}^{M} a_{j,m}\varpi_{j,m}^k, \qquad k = 1, \ldots, n; \qquad j = s, t. \tag{4.9}$$

where $\varpi_{j,m}^k = y_{j,m}^k / \overline{Y}_{j,m}$ is the share of $k^{\text{th}}$ DMU in the group in terms of the $m^{\text{th}}$-output (in period $j$). The corresponding price-independent weights for aggregating *allocative* efficiencies would then be

$$\tilde{S}_{j,ae}^k = \frac{S_j^k / D_\tau^k(x_j^k, y_j^k)}{\sum_{k=1}^{K} S_j^k / D_\tau^k(x_j^k, y_j^k)}, \qquad k = 1, \ldots, n; \quad j = s, t. \tag{4.10}$$

*Remark 5*. It is not the first time that 'positive' aggregation results in economics require some additional, often strong and perhaps undesirable assumptions (e.g., recall assumptions needed for aggregation of demands over consumers or over goods). In fact, in a more general context of aggregating efficiencies, Blackorby and Russell (1999) have proved *impossibility* results for their general case and the need of quite strong assumptions on technology in special cases.[7]

---

[6] If such information is unavailable, one might accept a more restrictive assumption: $a_{j,m}$ is the same constant for all $m$ (as in Färe and Zelenyuk (2003)).

[7] Our result does not contradict result of Blackorby and Russell (1999). Realizing that aggregation across *all* points in the technology set does not lead to a positive result, Färe and Zelenyuk (2003) established aggregation result for *optimal* points and use Mahler's inequality to obtain the aggregation result for the other points.

Application of (3.4), (4.2) and (4.3) immediately gives us the desired aggregation result (a solution to (3.5)) for the Malmquist Productivity Index, namely

$$\overline{RM}(p_s, p_t, \overline{Y_s}, \overline{Y_t}, X_s, X_t) =$$
$$= \left[ \left( \frac{\sum_{k=1}^{n} RE_s^k(x_t^k, y_t^k, p_t) \cdot S_t^k}{\sum_{k=1}^{n} RE_s^k(x_s^k, y_s^k, p_s) \cdot S_s^k} \times \frac{\sum_{k=1}^{n} RE_t^k(x_t^k, y_t^k, p_t) \cdot S_t^k}{\sum_{k=1}^{n} RE_t^k(x_s^k, y_s^k, p_s) \cdot S_s^k} \right)^{-1} \right]^{1/2}, \qquad (4.11)$$

and thus the desired decomposition for the *aggregate* level is

$$\overline{RM}(p_s, p_t, \overline{Y}_s, \overline{Y}_t, X_s, X_t) = \overline{M}(\cdot) \times \overline{AM}(\cdot), \qquad (4.12)$$

with the solutions to (3.7) and (3.8) given, respectively, by

$$\overline{M}(\overline{Y}_s, \overline{Y}_t, X_s, X_t) = \left[ \left( \frac{\overline{TE_s}(t)}{\overline{TE_s}(s)} \times \frac{\overline{TE_t}(t)}{\overline{TE_t}(s)} \right)^{-1} \right]^{1/2} \qquad (4.13)$$

where the four components inside (4.13) are given in (4.5), and

$$\overline{AM}(p_s, p_t, \overline{Y}_s, \overline{Y}_t, X_s, X_t) = \left[ \left( \frac{\overline{AE_s}(t)}{\overline{AE_s}(s)} \times \frac{\overline{AE_t}(t)}{\overline{AE_t}(s)} \right)^{-1} \right]^{1/2}. \qquad (4.14)$$

where the four components inside (4.14) are given in (4.6).

The theoretical and practical importance of these results is that we have obtained a way of aggregating the Malmquist Productivity Indexes—from the *individual* scores into *group* score. Importantly, in this approach, the aggregation function and, most importantly, the weights of aggregation are *not* ad hoc, but derived from economic principles (optimization) in such a way that the decomposition defined on the individual level is preserved at the group level.

We also would like to emphasize that we do not claim that the equally-weighted mean is useless for the context of Malmquist Productivity Indexes (or efficiency scores). On the opposite, we believe it should be used as a complementary descriptive statistic—as an estimator of the first moment of the distribution of the MPI (or efficiency scores), if it exists. We argue however, that it must be supported and compared with the average that accounts for some economic weight of each observation.

Our derivations were performed for the output orientation case. Similar developments are easily transferred for input orientation, where cost minimization would be used as a criterion to derive the weights for input oriented indexes. This would be an intertemporal extension of Färe, Grosskopf and Zelenyuk (2002a). Similar analysis can also be done for aggregation of the productivity indexes defined in terms of the directional distance functions (this would be an intertemporal extension of the results found in Färe, Grosskopf and Zelenyuk (2002a, 2002b)).

## 5. Geometric vs. Harmonic Averaging of Malmquist Productivity Indexes

As was mentioned above, in practice, to summarize a large number of estimated Malmquist productivity indexes in a single number, researchers often resorted to equally-weighted averages of the *individual* estimates. Moreover, the aggregation function was the *geometric* average—a tradition that started at least with the seminal paper of Färe et al. (1994), and was motivated by the multiplicative nature of the index (see their footnote 19 on p. 78; also see Färe and Zelenyuk, 2002, for related formal discussions). In the previous section, our derivations have given us both: the system of weight and in the aggregation function. Clearly, the weights might be critically important for drawing both quantitative and qualitative conclusions. The question of functional form of aggregation is not that clear at this stage.

The goal of this section is to investigate the relationship (or/and the difference) between the aggregate efficiency measure based on the geometric aggregation and the harmonic one that we have derived above, assuming both use the *same* system of weights. To establish the relationship between these two aggregation approaches, we first use (4.5) to rewrite $\overline{M}(\cdot)$ in (4.13) in terms of the four components of *harmonic* aggregations of scores from individual distance functions,

$$\overline{M}(\cdot)=\left[\left(\frac{\left(\sum_{k=1}^{n}[D_s^k(x_t^k,y_t^k)]^{-1}\cdot S_t^k\right)^{-1}}{\left(\sum_{k=1}^{n}[D_s^k(x_s^k,y_s^k)]^{-1}\cdot S_s^k\right)^{-1}}\times\frac{\left(\sum_{k=1}^{n}[D_t^k(x_t^k,y_t^k)]^{-1}\cdot S_t^k\right)^{-1}}{\left(\sum_{k=1}^{n}[D_t^k(x_s^k,y_s^k)]^{-1}\cdot S_s^k\right)^{-1}}\right)\right]^{1/2}. \tag{5.1}$$

In general, the geometric analogue of (5.1) is defined (for some weights $W_t^k, W_s^k$) as

$$\overline{M}^G(\cdot)\equiv\left[\frac{\prod_{k=1}^{n} D_s^k(x_t^k,y_t^k)^{W_t^k}}{\prod_{k=1}^{n} D_s^k(x_s^k,y_s^k)^{W_s^k}}\times\frac{\prod_{k=1}^{n} D_t^k(x_t^k,y_t^k)^{W_t^k}}{\prod_{k=1}^{n} D_t^k(x_s^k,y_s^k)^{W_s^k}}\right]^{1/2}, \tag{5.2}$$

(The aggregation that is commonly used in practice is a special case of (5.2) that assumes equal weights across all $k$.) Clearly, the formulation in (5.1) yields, in general, different values than those from (5.2), and there is no *exact* general relationship between them. However, if we look at the first-order approximation of $\prod_{k=1}^{n} D_s^k(x_t^k,y_t^k)^{S_t^k}$ and of $\left(\sum_{k=1}^{n}[D_s^k(x_t^k,y_t^k)]^{-1}\cdot S_t^k\right)^{-1}$ around *unity*—which is a natural point around which an approximation of productivity and efficiency indexes can be done (see Färe and Zelenyuk, 2002), then they both are equal to $\sum_{k=1}^{n} D_s^k(x_t^k,y_t^k)\cdot S_t^k$. This leads us to a desired relationship implying that

$$\overline{M}(Y_s,Y_t,X_s,X_t)\cong\overline{M}^G(Y_s,Y_t,X_s,X_t),\qquad \text{for } (W_t^k,W_s^k)=(S_t^k,S_s^k). \tag{5.3}$$

In other words, expression (5.3) establishes the *first-order-approximation* relationship between the *harmonic* aggregate Malmquist Productivity Index we have derived in previous sections and the commonly used *geometric* aggregate of individual MPI's—given that both aggregations use the same system of weights. Thus, for a researcher who prefers the geometric aggregation (e.g., for multiplicativity reasons) this relationship allows justifying the choice of weights of aggregation—the weights derived from economic optimization that attempt accounting for economic importance of each DMU in the sample.

A natural question is how different the results of the geometric and harmonic aggregations would be in practice. To get some feeling on this, we provide some of the typical results of Monte Carlo experiments we had. Here we present only 7 scenarios, with $n$ = 100, and with 1000 replications in each Monte Carlo experiment. The basic scenario assumes that the scores of Malmquist productivity indexes are distributed uniformly between 0.5 and 1.5. This allows for quite some variation in the scores: the range is 100 percentage points of productivity change, and standard deviation is about 29 percentage points of productivity change, which is more than what is often observed in empirical studies analyzing short-run changes in productivity. The other scenarios consider uniform distributions for ranges of (0.3, 1.7), (0.1,1), (1, 2.5), (0.5,1), (1,1.5), (0.75,1.25). The (non-equal) weighting scheme is the same for both types of aggregation and for each scenario was generated from the *uniform* distribution on (0,1) and then normalized to sum to one.

**Table 1.** Difference between the Geometric and Harmonic Aggregations of MPI: Monte Carlo Results for Four Simulated Scenarios

| *Ranges of Uniform* | $\sqrt{\sum_{b=1}^{B}(d_b)^2} / B$ | $\sum_{b=1}^{B} \mid d_b \mid / B$ | $\min_b\{d_b\}$ | $\max_b\{d_b\}$ |
|---|---|---|---|---|
| (0.5, 1.5) | 0.0452 | 0.0450 | -0.0613 | -0.0320 |
| (0.3, 1.7) | 0.0989 | 0.0983 | -0.1374 | -0.0622 |
| (0.1, 1) | 0.0839 | 0.0833 | -0.1181 | -0.0508 |
| (1, 2.5) | 0.0562 | 0.0559 | -0.0743 | -0.0405 |
| (0.5, 1) | 0.0143 | 0.0142 | -0.0191 | -0.0094 |
| (1, 1.5) | 0.0084 | 0.0083 | -0.0108 | -0.0056 |
| (0.75, 1.25) | 0.0105 | 0.0105 | -0.0142 | -0.0072 |

*Notes:* $d_b$ = difference between the Harmonic and Geometric aggregations of simulated values for distance functions in Monte Carlo replication $b$ ($b$ =1,…, $B$ = 1000); $n$ = *100*.

Results in Table 1 suggest that, as expected, the geometric aggregation always gives an upward biased aggregate productivity score relative to the harmonic aggregation (assuming the same weighting scheme for the two aggregations). The difference however is not very large. Even when the range of the distribution of productivity changes is 100 percentage points—the square root of the mean square difference (SRMSD) of the two aggregate indexes is 4.5 percentage points (with the maximum being 6.1), over 1000 replications. When the range is 50 percentage points, then only about 1 percentage point of the SRMSD is observed. For scenarios where the mean is not unity (i.e., zero change, around which we approximate), but is within 25 percentage points of it, then the SRMSD is still only about 1 percentage point.

A practical conclusion we can draw from this subsection is that the proposed harmonic-type and geometric-type aggregations of the productivity indexes, whose first-order approximations are equivalent under the same weighting scheme, give very similar aggregate scores for distributions with a quite wide spread of individual scores. This may justify the use of geometric aggregations, if preferred. However we suggest using the economically justified weighting scheme derived above—which clearly may give very different results from the commonly used equally-weighted aggregation.

## 6. Decomposition into the Aggregate Technical and Aggregate Efficiency Changes

The aggregation results derived in sections 4 and 5 above can easily be extended to the aggregation of components of various conceptual decompositions of Malmquist Productivity Indexes (see for example Balk (2004) for a recent survey). Here, for the sake of brevity, we limit ourselves to only one of the most popular decompositions suggested by Färe et al. (1994), defined as

$$M^k(\cdot) \equiv EFCH^k(\cdot) \times TECH^k(\cdot), \tag{6.1}$$

where *change in efficiency* is measured by

$$EFCH^k(\cdot) \equiv EFCH^k(y_s^k, y_t^k, x_s^k, x_t^k) \equiv \frac{D_t^k(x_t^k, y_t^k)}{D_s^k(x_s^k, y_s^k)}, \tag{6.2}$$

and *technological change* is measured by

$$TECH^k(\cdot) \equiv TECH^k(y_s^k, y_t^k, x_s^k, x_t^k) \equiv \left[\frac{D_s^k(x_t^k, y_t^k)}{D_t^k(x_t^k, y_t^k)} \times \frac{D_s^k(x_s^k, y_s^k)}{D_t^k(x_s^k, y_s^k)}\right]^{1/2}. \tag{6.3}$$

We now want our *group* analogs to be obtainable from aggregating the corresponding individual measures (6.2) and (6.3) or its components, given by (2.2), via some function $f_E(\cdot)$, and $f_{TC}(\cdot)$. Given our developments above, a natural choice would be to set

$$\overline{EFCH}(\cdot) = \left(\frac{\overline{\overline{TE}_t(t)}}{\overline{\overline{TE}_s(s)}}\right)^{-1} = \frac{\left(\sum_{k=1}^n [D_t^k(x_t^k, y_t^k)]^{-1} \cdot S_t^k\right)^{-1}}{\left(\sum_{k=1}^n [D_s^k(x_s^k, y_s^k)]^{-1} \cdot S_s^k\right)^{-1}}, \tag{6.4}$$

and

$$\overline{TECH}(\cdot) = \left[\left(\frac{\overline{\overline{TE}_s(t)}}{\overline{\overline{TE}_t(t)}} \times \frac{\overline{\overline{TE}_s(s)}}{\overline{\overline{TE}_t(s)}}\right)^{-1}\right]^{1/2}$$

$$= \left[\frac{\left(\sum_{k=1}^n [D_s^k(x_t^k, y_t^k)]^{-1} \cdot S_t^k\right)^{-1}}{\left(\sum_{k=1}^n [D_t^k(x_t^k, y_t^k)]^{-1} \cdot S_t^k\right)^{-1}} \times \frac{\left(\sum_{k=1}^n [D_s^k(x_s^k, y_s^k)]^{-1} \cdot S_s^k\right)^{-1}}{\left(\sum_{k=1}^n [D_t^k(x_s^k, y_s^k)]^{-1} \cdot S_s^k\right)^{-1}}\right]^{1/2} \tag{6.5}$$

Similar to the previous section, the first order approximation relationship exist between the harmonic-type aggregations (6.4) and (6.5) and their geometric analogues.

This completes our brief outline of the main results that can be used by *practitioners* for summarizing their estimation results of individual Malmquist Productivity Indexes into *group* (or aggregate) Malmquist Productivity Indexes which attempt accounting for the economic importance of each observation in the sample via a theoretically justified weighting scheme.

## 7. Conclusion

In this paper we have extended the work of Färe and Zelenyuk (2003) to obtain a theoretically justified method for aggregating Malmquist Productivity Indexes and their decompositions. We do not suggest that the new aggregate measures and their aggregate components of decompositions must replace the commonly used equally-weighted analogs. Instead, we suggest that they shed important *additional* light in the analysis of productivity changes—since they attempt accounting for an economic importance of each observation in the sample—not in an ad hoc way, but using weights derived from economic optimization criterion. Also noteworthy is that the same aggregation principle can be applied to obtain aggregation results for other indexes that are based on revenue, cost, profit, directional and Shephard's distance functions (for example, for aggregation of price indexes and Malmquist quantity indexes, Hicks-Moorsteen indexes, etc).

A natural further extension would be to develop methods of statistical inference on the group (and sub-groups) Malmquist (and other) productivity indexes, which can be done, for example, by merging the ideas of Simar and Wilson (1999) with Simar and Zelenyuk (2003).

## References


Blackorby, C., and R. Russell (1999), "Aggregation of Efficiency Indices", *Journal of Productivity Analysis* 12:1, 5-20.

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

Kuosmanen, T., L. Cherchye, and T. Sipiläinen (2004) "The Law of One Price in Data Envelopment Analysis: Restricting Weight Flexibility Across Firms," paper presented at NAPW2004 (Toronto, Canada), *forthcoming* in *European Journal of Operational Research*.

Balk, B., (2004) "The Many Decompositions of Productivity Change," paper presented at NAPW2004 (Toronto, Canada).

Domar, E. (1961), "On the Measurement of Technological Change," *The Economic Journal* 71:284, pp.709-729.

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

Färe, R., S. Grosskopf and V. Zelenyuk (2002a) "Aggregation of Cost Efficiency Indicators and Indexes Across Firms", *forthcoming* in *Academia Economic Papers*.

Färe, R., S. Grosskopf and V. Zelenyuk (2002b) "Aggregation of the Nerlovian Profit Indicator," Mimeo (submitted).

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

Färe, R. and V. Zelenyuk (2003), "On Aggregate Farrell Efficiency Scores," *European Journal of Operational Research* 146:3, 615-620.

Färe, R. and V. Zelenyuk (2002), "Averaging Farrell Scores," Mimeo (submitted).

Farrell, M.J. (1957), "The Measurement of Productive Efficiency," *Journal of Royal Statistical Society*, Series A, General, 120, part 3, 253-281.

Henderson, D. and V. Zelenyuk (2003) "Testing for Catching-up: Statistical Analysis of DEA Efficiency Estimates," Disc.Paper 0431 of Institute of Statistics, University Catholique de Louvain.

Koopmans, T.C. (1957), *Three Essays on the State of Economic Analysis*, New York: McGraw-Hill.

Russell, R.R., (1990), "Continuity of Measures of Technical Efficiency," *Journal of Economic Theory* 51, 255-267.

Russell, R.R., (1997), "Distance Functions in Consumer and Producer Theory," in *Index Number Theory: Essays in Honor of Sten Malmquist*, Kluwer Academic Publishers, 7-90.

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

Simar, L. and P. Wilson (1999), " Estimating and Bootstrapping Malmquist Indexes," *European Journal of Operational Research* 115, 459-471.

Simar, L. and V. Zelenyuk (2003), "Statistical Inference for Aggregates of Farrell-type Efficiencies," *Discussion Papers (0324) of Institute of Statistics, University Catholique de Louvain*, Belgium.