



Centre for Efficiency and Productivity Analysis

**Working Paper Series
No. WP06/2019**

Aggregation of Efficiency and Productivity:

From Firm to Sector and Higher Levels

Valentin Zelenyuk

Date: May 2019

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

ISSN No. 1932 - 4398

Aggregation of Efficiency and Productivity: From Firm to Sector and Higher Levels

Valentin Zelenyuk*

May 3, 2019

Abstract

Here we consider various cases where researchers are interested in measuring *aggregate* efficiency or productivity levels or their changes for a *group* of decision making units. These could be entire industry composed of individual firms, banks, hospitals, or a region composed of sub-regions or countries, or particular sub-groups of these units within a group, e.g., sub-groups of public vs. private or regulated vs. non-regulated firms, banks or hospitals within the same industry, etc. Such analysis requires solutions to the aggregation problem—some theoretically justified approaches that can connect individual measures to aggregate measures. Various solutions are offered in the literature and our goal is to try to coherently summarize at least some of them in this chapter. This material should be interesting not only for theorists but also (and perhaps more so) for applied researchers, as it provides exact formulas and intuitive explanations for various measures of group efficiency, group scale elasticity and group productivity indexes and refers to original papers for more details.

Keywords: Efficiency, Productivity, Aggregation, Industry Efficiency, Duality.

JEL Classification: D24, C43, L25

*School of Economics and Centre for Efficiency and Productivity Analysis (CEPA) at The University of Queensland, Australia; *address:* 530, Colin Clark Building (39), St Lucia, Brisbane, Qld 4072, AUSTRALIA; e-mail: v.zelenyuk@uq.edu.au; tel: + 61 7 3346 7054; The author acknowledges support of the University of Queensland and from the ARC grants (ARC FT170100401).

1 Introduction

An aggregate perspective is very important for theory and perhaps even more so in practice. Even if a researcher estimates the efficiency of individual units, she/he might still (and usually do) want to have just one or a few aggregate numbers that summarize the individual estimates. Such aggregate numbers would be especially useful if the number of individual units is too large to report all of them, and especially to comprehend them all for understanding the overall picture. Indeed, hardly anyone would want to read hundreds of individual efficiency scores and would rather demand a summary—some *aggregate* efficiency or productivity measures that will give a big picture about the efficiency or productivity situation in the industry or sub-groups of interest within it.

The key question here is therefore: How to meaningfully aggregate the individual efficiency and productivity scores or indexes? A natural answer would be: ‘Take an average!’ But, which one? Is it arithmetic, geometric, harmonic, or any other? And, much more importantly: Should it be a weighted or a non-weighted average? Or, more generally: What are the most meaningful (from economics point of view) ways to aggregate the individual efficiency and productivity scores or indexes, of potentially many individuals, into one number representing aggregate efficiency or productivity of a group?

The goal of this chapter is to provide some answers to these fundamentally important questions, by summarizing the recent developments in the literature. In a nutshell, the results summarized here provide applied researchers with the formulas for group efficiency, group scale elasticity and group productivity indexes. Importantly, in all these formulas, the weights of aggregation have a fairly intuitive economic meaning, yet they are not *ad hoc* but derived mathematically via economic theoretical reasoning.¹

2 The Aggregation Problem: A Brief Background

The problem of finding a measure (a score, an index) representing a group of individual measures is called an aggregation problem—a problem that has been studied in many fields, including economics.

In the field of productivity and efficiency analysis, this problem have been raised starting, at least, with the classical works of Farrell (1957), Førsund and Hjalmarsson (1979), and

¹A different area of the aggregation questions that focuses on the aggregation of inputs or aggregation of outputs for a firm (e.g., to reduce the dimension of the model) is not considered here and can be found in Tauer (2001); Färe and Zelenyuk (2002); Färe et al. (2004a); Wilson (2018) and the references therein. We also do not consider the question of aggregation of indexes with respect to different references (e.g., time periods) for the same firm, which can be found in Färe and Zelenyuk (2019) and the references therein.

later followed up by Li and Ng (1995), and most thoroughly theoretically scrutinized by Blackorby and Russell (1999) and critically evaluated by Ylvinger (2000).

More recently this important analytical problem was addressed and to some extent resolved by Färe and Zelenyuk (2003); Färe et al. (2004b), Färe and Zelenyuk (2005, 2007), Simar and Zelenyuk (2007), Färe et al. (2004b), Zelenyuk (2006), Simar and Zelenyuk (2007), Nesterenko and Zelenyuk (2007) and most recently by Mayer and Zelenyuk (2014a, 2017), to mention a few. Here, we will briefly summarize the essence of the key results from these and other works.²

2.1 The Essence of the Aggregation Problem

As in the general context, the most important issue here is the choice of *weights* in the aggregation. To vividly illustrate the point, consider an example of an industry with many firms, most of which are small, while a very few large firms take most of the industry share.³ Now suppose that those small firms are very efficient and suppose for simplicity of computation they are 100% (or nearly that) efficient. Meanwhile, suppose those big firms are much less efficient, say 50% efficient.⁴ If for such an example, a researcher were to use the simple (equally-weighted) arithmetic average then such aggregate efficiency score would indicate the industry is about close to 100% efficient! On the other hand, if another researcher wanted to use a *weighted* arithmetic average, then a dramatically different conclusion might be reached—depending on the weighting scheme. Indeed, if one takes the market shares as the relative weights and uses them in the weighted arithmetic average, then such an aggregate efficiency score will indicate that the industry is closer to 50% efficient. Thus, one would reach a dramatically different conclusion with opposite policy implications than from the equally-weighted average!

The essence of the problem here is in the nature of efficiency scores—by construction, they are ‘standardized’ so that they are between 0 and 1 and, while this gives some advantages, the side-effect of such standardization is that they lose the information about the relative weights of the firms that obtained these scores. Clearly, one may try to justify some other

²This chapter is a substantially revised, extended and elaborated material that I presented earlier, in Chapter 5 of Sickles and Zelenyuk (2019).

³While this is a generic example, a reader might have realized that many industries in the real world have a similar composition, often resembling the so-called ‘Pareto principle’, more casually known as ‘the 80/20 rule’ postulating that about 80% share (e.g., of wealth, sales, etc.) is taken up by about 20% of members of a group.

⁴Lower efficiency of large firms is not unusual and often was reported in the literature. It can arise, for example, due to the greater complexity of being a larger organization involving greater levels of hierarchy and thus implying potentially greater principal-agent problems or requiring more inputs or higher costs than needed for producing the same level and the same quality of output.

weights that may imply very different conclusions and thus different policy implications and this, in turn, emphasizes the importance of having justifications for the choice of weights.

2.2 The Evolution of the Aggregation Literature

The early key ideas that attempted to take into account the economic weights of firms when aggregating their efficiency can be found in the seminal work of Farrell (1957), where he proposed the concept of ‘*Structural Efficiency of an Industry*’. To be precise, Farrell (1957) considered a single-output case and proposed taking the *weighted arithmetic average* of efficiency scores of individual firms in that industry, where the *weights* were the observed output shares of the firms within the industry. Importantly, note that Farrell had not given any formal theoretic justifications for such an aggregation scheme at that time and, in particular, had not justified why *output* shares were to be used for aggregating the *input* oriented technical efficiency scores that he considered.⁵ Farrell also did not explain how to apply his idea for a multiple-output case. These limitations were perhaps among the main reasons for why Farrell’s concept of Structural Efficiency of an Industry had not attained wider use in practice, where many tended to just use the equally weighted averages to report on the aggregate efficiency of industries or sub-groups within them.

About two decades later, Farrell’s ideas were revisited by Førsund and Hjalmarsson (1979) who proposed to estimate efficiency scores of an ‘average firm’, constructed as the average of input-output allocations. It is easy to construct an example that will show that such a measure can indicate high inefficiency even though all firms are technically efficient and this was considered (incorrectly) as a drawback and, apparently was one of the reasons why this measure is also rarely used in practice.

The fundamental ideas of Farrell (1957) and Førsund and Hjalmarsson (1979) were then revisited by a very important (yet for a long time overlooked) work of Li and Ng (1995), who attempted to synthesize the two approaches with additional assumptions, focusing on the data envelopment analysis (DEA) context, and on the use of the so-called shadow prices in DEA.

At the turn of the last century, Blackorby and Russell (1999) were the first to scrutinize the problem on pure theoretical grounds and derived several important, yet ‘negative’ results—they proved the impossibility of a solution of the aggregation problem in a general setup. What this implied was that some additional assumptions or structure were needed to arrive at a ‘positive result’.⁶ Such additional assumptions and structure were discovered by

⁵Indeed, later in this chapter we will see that output shares are more coherent with output orientation, while for the input orientation it would be more natural to use the cost shares.

⁶This is not entirely surprising, e.g., recall that very strong assumptions are needed to establish positive

Färe and Zelenyuk (2003): In addition to the usual assumptions of production theory, they followed Koopmans (1957) work on aggregation in economics, adapting it to the context of efficiency analysis. Differently from Koopmans (1957), however, they assumed an additive structure for the aggregate technology being the set-wise summation of the individual output (rather than technology) sets, for given input allocations. Adding this structure to the standard regularity conditions of production theory and with the so-called ‘law of one price’ assumption (as in Koopmans), Färe and Zelenyuk (2003) then involved the principles of *economic optimization* to derive a theoretically justified weighting scheme for aggregation of individual efficiencies into a group efficiency. In turn, this theoretical framework provided the grounds of economic theory for the weighting scheme of Farrell (1957) and Li and Ng (1995) and circumvented the impossibility theorems of Blackorby and Russell (1999).

The approach of Färe and Zelenyuk (2003) was then used to derive many other interesting and useful aggregation results: e.g., for aggregation of directional distance functions in Färe et al. (2008), aggregation of scale elasticities in Färe and Zelenyuk (2012) and scale efficiencies in Zelenyuk (2015), for aggregation of Malmquist and Hicks-Moorsteen productivity indexes Zelenyuk (2006); Mayer and Zelenyuk (2014a, 2017), etc. The goal of this chapter is to summarize these aggregation results and give some insights on future developments.

3 Aggregation of Efficiency Scores

While most of the discussion here will be theoretical, it would be helpful for a reader to keep in mind that a typical empirical context of this methodology is a study of efficiency or productivity (or their changes) of an economic system consisting of different decision making units (hereafter DMUs), e.g., industry consisting of firms or a particular bank or other institutions consisting of many branches, or a comparison of *distinct groups* within such a system (e.g., regulated vs. non-regulated, foreign vs. local, private vs. public firms, etc.). To get to the group level, we first need to briefly refresh the key concepts and notions for the individual level, which we do in the next sub-section.

3.1 Individual Primal and Dual Efficiency Scores

Without loss of generality, suppose the system is a group (e.g., industry, sector, etc.) consisting of n DMUs, where for each DMU $k \in \{1, 2, \dots, n\}$ we will use vector $x^k = (x_1^k, \dots, x_N^k)' \in \mathfrak{R}_+^N$ to denote N *inputs* that the DMU k utilizes to produce a vector of M *outputs*, $y^k = (y_1^k, \dots, y_M^k)' \in \mathfrak{R}_+^M$. For generality of the aggregation results, we will allow for each DMU k

aggregation results in consumer theory.

to employ technology that is potentially different from those used by other DMUs, and we assume it can be characterized by the technology set Ψ^k , defined in general terms as⁷

$$\Psi^k \equiv \{(x^k, y^k) \quad : \quad x^k \text{ can produce } y^k\}. \quad (3.1)$$

An equivalent characterization of technology can also be given via the *output sets*

$$P^k(x^k) \equiv \{y^k \quad : \quad x^k \text{ can produce } y^k\}, \quad x^k \in \mathfrak{R}_+^N. \quad (3.2)$$

An important advantage of the aggregation results that we summarize here is their generality with respect to characterization of technology. Indeed, we do not assume any particular production or transformation function (e.g., Cobb-Douglass, Leontieff, CES), rather we allow for a very wide class of technologies that satisfy usual regularity axioms of production theory, and in particular:

A1: The technology set Ψ^k is closed.

A2: The output correspondence $P^k(x^k)$ is bounded $\forall x^k \in \mathfrak{R}_+^N$.

A3: There is no ‘free lunch’, i.e. nothing cannot produce something, i.e.,

$(0_N, y^k) \notin \Psi^k, \quad \forall y^k \geq 0_M$ (i.e. $y_m^k \geq 0$ for $m = 1, \dots, M, y^k \neq 0_M$).

A4: It is possible to produce nothing, i.e., $0_M \in P^k(x^k), \quad \forall x^k \in \mathfrak{R}_+^N$.

A5: Outputs and inputs are freely (strongly) disposable, i.e.,

$(x^0, y^0) \in \Psi^k \implies (x, y) \in \Psi^k, \quad \forall y \leq y^0, \forall x \geq x^0$.

To employ the results from the duality theory in economics we also need some convexity assumptions. At the beginning we only assume that the output sets are convex, i.e.,

A6: $y^0, y^1 \in P^k(x^k), \quad x^k \in \mathfrak{R}_+^N \implies \delta y^0 + (1 - \delta)y^1 \in P^k(x^k), \quad \forall \delta \in [0, 1]$.⁸

With these conditions, the *output oriented* Shephard (1970) distance function $D_o^k : \mathfrak{R}_+^N \times \mathfrak{R}_+^M \rightarrow \mathfrak{R}_+^1 \cup \{+\infty\}$, defined as

$$D_o^k(x^k, y^k) \equiv \inf_{\theta} \{\theta > 0 : y^k / \theta \in P^k(x^k)\}, \quad (3.3)$$

gives a complete characterization of the technology of a DMU k , in the sense that

$$D_o^k(x^k, y^k) \leq 1 \quad \Leftrightarrow \quad y^k \in P^k(x^k). \quad (3.4)$$

A closely related concept is the *Farrell* output oriented measure of *technical efficiency*,

⁷In the discussion of economic theoretical foundation here we mainly use framework developed by Shephard (1953, 1970) and further refined in many works and concisely outlined in Färe and Primont (1995) and Sickles and Zelenyuk (2019).

⁸For theoretical results we do not require convexity of Ψ^k , although when implementing in practice one may impose it when choosing a particular estimator or particular functional form for technology.

defined as

$$OTE^k(x^k, y^k) \equiv \sup_{\theta} \{\theta > 0 : \theta y^k \in P^k(x^k)\} = 1/D_o^k(x^k, y^k). \quad (3.5)$$

Furthermore, let $p = (p_1, \dots, p_M) \in \mathfrak{R}_{++}^M$ be the vector of corresponding output prices⁹ then the *dual* characterization of $P^k(x^k)$ is obtained from the *revenue* function, $R^k : \mathfrak{R}_+^N \times \mathfrak{R}_{++}^M \rightarrow \mathfrak{R}_+^1 \cup \{+\infty\}$

$$R^k(x^k, p) \equiv \sup_y \{py \quad : \quad y \in P^k(x^k)\}, \quad (3.6)$$

and the related efficiency measure for a DMU k in the *dual* framework would then be the *revenue efficiency* (also referred to as or the overall output efficiency), defined formally as,

$$RE^k(x^k, y^k, p) \equiv R^k(x^k, p)/py^k. \quad (3.7)$$

From the duality theory for the revenue function (Shephard (1970); Färe and Primont (1995)), we then have

$$R^k(x^k, p) \geq py^k/D_o^k(x^k, y^k), \quad (3.8)$$

which leads to another notion—a measure of the output oriented *allocative* (in)efficiency, defined as a multiplicative residual that turns (3.8) into equality, i.e.,

$$OAE^k(x^k, y^k, p) \equiv RE^k(x^k, y^k, p)/OTE^k(x^k, y^k), \quad (3.9)$$

and so, we have a useful decomposition:

$$RE^k(x^k, y^k, p) = OTE^k(x^k, y^k) \times OAE^k(x^k, y^k, p), \quad (3.10)$$

This decomposition (3.10) is a stepping-stone for deriving the aggregation results, as will be apparent below.

3.2 Group Primal and Dual Efficiency Scores

Let us consider a *sub-group* l ($l = 1, \dots, L$), consisting of n_l DMUs within the original group of n DMUs. Such sub-grouping can be based on various exogenous criteria such as geographic regions, ownership structures, regulation regimes, etc. For each group l ($l = 1, \dots, L$), let the input allocation among DMUs *within* the sub-group l be $X^l = (x^{l,1}, \dots, x^{l,n_l})$ and let the *total* of output vectors over all firms in the l^{th} group be $\bar{Y}^l = \sum_{k=1}^{n_l} y^k$.

A cornerstone in the derivation of the aggregation results is the structure of the *aggregate*

⁹Note that for the aggregation results, a necessary assumption is the so-called ‘Law of One Price’, i.e., here it implies that all firms face the same output prices.

technology. In the context of output orientation, it is natural to assume a linear structure of aggregation of the output sets, as was done in Färe and Zelenyuk (2003): For each group l ($l = 1, \dots, L$), the aggregate output set $\bar{P}^l(X^l)$ is the Minkowski sum of the individual output sets across all DMUs k ($k = 1, \dots, n_l$) within the group l , i.e.,¹⁰

$$\bar{P}^l(X^l) \equiv \sum_{\oplus k=1}^{n_l} P^{l,k}(x^{l,k}). \quad (3.11)$$

As a result of such a structure, $\bar{P}^l(X^l)$ would inherit the regularity conditions imposed on the individual output sets. In particular, note that the Minkowski sum of convex sets is also a convex set.¹¹ Thus, convexity of the individual output sets imposed by A6 ensures convexity of $\bar{P}^l(X^l)$.

It is also worth noting that the aggregation structure defined by $\bar{P}^l(X^l)$ presumes no reallocation of inputs across the individuals $k \in \{1, \dots, n_l\}$ and so depends not on the total sum of all the inputs but on the particular allocation $X^l = (x^{l,1}, \dots, x^{l,n_l})$. This structure also assumes there are no externalities across firms.

Now, using the l^{th} sub-group technology (3.11), one can define the *sub-group revenue* function as

$$\bar{R}^l(X^l, p) \equiv \sup_y \{py : y \in \bar{P}^l(X^l)\}, \quad (3.12)$$

which, analogously to (3.7), gives rise to the l^{th} sub-group *revenue efficiency* measure

$$\overline{RE}^l(X^l, \bar{Y}^l, p) \equiv \bar{R}^l(X^l, p) / p\bar{Y}^l. \quad (3.13)$$

3.3 The Fundamental Aggregation Results

Having specific formulas for the efficiency measures defined with respect to individual technologies and with respect to aggregate technologies raises questions regarding the relationship between them. Ideally, one may want to establish their equality, so that the latter can be obtained from the former via some feasible computations, at least under some clear and reasonable conditions. Formally, the goal is to find $f_{RE}(\cdot)$ such that

$$\overline{RE}^l(X^l, \bar{Y}^l, p) = f_{RE}(RE^1(\cdot), \dots, RE^{n_l}(\cdot)). \quad (3.14)$$

¹⁰We use \oplus to distinguish the summation of sets (also called ‘Minkowski summation’) from the standard summation, e.g., see Oks and Sharir (2006).

¹¹E.g., see Krein and Smulian (1940), Schneider (1993), and a more recent work of Oks and Sharir (2006), as well as references therein. For other examples involving Minkowski summation in economics, see Shapley–Folkman–Starr theorem and related results (Starr (2008)).

In words, the goal is to find some aggregation function, which we call $f_{RE}(\cdot)$, that can relate the aggregate measure (3.13) to the individual measures (3.7), for all firms $k \in \{1, \dots, n_l\}$ in a group of interest, and do so in some meaningful way in the sense that the group measure should represent the group. Finding a function $f_{RE}(\cdot)$ is not a difficult problem—there is an abundance of well-studied functions offered by mathematicians. It is the ‘meaningful way’ aspect that is the most challenging and, as with many (if not all) notions in economics, depends on the views and assumptions of a researcher. The goal therefore is to make the choice grounded on and derived from some clear assumptions and if one does not like some assumptions then one may try to replace them with others and, possibly, derive new aggregation results. This is the approach we discuss here. In particular, we also consider it as desirable that the decomposition of revenue efficiency into technical efficiency and allocative efficiency that we have at the individual level is also maintained at the *aggregate* level, so that we have

$$\overline{RE}^l(\cdot) = \overline{OTE}^l(\cdot) \times \overline{OAE}^l(\cdot), \quad (3.15)$$

where

$$\overline{OTE}^l(\cdot) = f_{TE}(TE^1(\cdot), \dots, TE^{n_l}(\cdot)), \quad (3.16)$$

and

$$\overline{OAE}^l(\cdot) = f_{AE}(OAE^1(\cdot), \dots, OAE^{n_l}(\cdot)), \quad (3.17)$$

where $f_{TE}(\cdot)$, $f_{AE}(\cdot)$ are also some aggregation functions to be found (potentially different from each other and from $f_{RE}(\cdot)$) so that they ensure the aggregate measures are related to the individual analogues. Such functions can be found using the following fundamental theorem.

Theorem 1. *For each group l ($l = 1, \dots, L$), the maximal revenue of the sub-group of DMUs feasible from (X^l, p) is equal to the sum of the maximal revenues of all its member DMUs feasible from their $(x^{l,k}, p)$, $k = 1, \dots, n_l$, i.e.,*

$$\bar{R}^l(X^l, p) = \sum_{k=1}^{n_l} R^{l,k}(x^{l,k}, p). \quad (3.18)$$

This theorem is from Färe and Zelenyuk (2003) and it can be viewed as the revenue analog to the Koopmans (1957) theorem of aggregation of the profit functions, while the cost or input-oriented analog can be found in Färe et al. (2004b)).

More importantly, this theorem provides a key to our aggregation problem and so it is important to understand the economic intuition behind it: The theorem says that the sum of the revenues of individual revenue-maximizing DMUs in a sub-group is the same as the

revenue optimized over the aggregate technology (3.11) for this sub-group, provided these DMUs face the same (e.g., equilibrium) output prices (and other regularity conditions hold). That is, whether optimized individually or as a group, the same revenue is attained under the ‘Law of One Price’ (e.g., equilibrium price level) for all the outputs.

The theorem above assumes full revenue efficiency (and full information) and so, a natural question is: Why do we consider full revenue efficiency when we want to measure output oriented inefficiency? And the answer is: Because we need it to set a benchmark against which the inefficiency will be measured. This is in the same fashion as how we choose the maximal output as the benchmark (although not assuming it to be reached by each firm) so that the actual output can be measured relative to it, in the output oriented context of efficiency measurement.¹²

From this fundamental theorem (as well as its cost and profit analogues) one can then get many useful results for the aggregation of the efficiency scores, some of which we summarize below, starting with the following corollary that first appeared in Färe and Zelenyuk (2003) and is an immediate consequence of (3.18).

Corollary 1. *For each group l ($l = 1, \dots, L$), we have*

$$\overline{RE}^l(X^l, \overline{Y}^l, p) = \sum_{k=1}^{n_l} RE^{l,k}(x^{l,k}, y^{l,k}, p) \times S^{l,k}, \quad (3.19)$$

where

$$S^{l,k} = py^{l,k}/p\overline{Y}^l, \quad k = 1, \dots, n_l. \quad (3.20)$$

In words, this corollary states that the weighted sum of the revenue efficiencies of individual revenue-maximizing firms in a sub-group is the same as the revenue efficiency with respect to the aggregate technology (3.11) for this same sub-group, provided these firms face the same output prices and the standard regularity conditions of production theory hold. In turn, this corollary implies another useful result, which gives the weighting schemes for the technical and allocative efficiencies into their group analogues, preserving the decomposition like (3.10) also at the aggregate level. We summarize this important result in the next corollary (also first appeared in Färe and Zelenyuk (2003)).

For each group l ($l = 1, \dots, L$), the aggregate revenue efficiency can be decomposed multiplicatively into the weighted sum of the technical efficiencies (where the weights are

¹²In the input oriented context, such a benchmark will be the cost function, while in the framework where both input and output vectors can be changed when measuring efficiency (e.g., for efficiency based on the directional distance function or hyperbolic measures), the natural benchmark will be the profit function. We will briefly discuss these cases later in the chapter.

the actual revenue shares) and the weighted sum of the allocative efficiencies (where the weights are the revenue shares corrected for technical inefficiency) of all its member DMUs. We summarize this formally in the next corollary.

Corollary 2. *For each group l ($l = 1, \dots, L$), we have*

$$\overline{RE}^l(X^l, \bar{Y}^l, p) = \overline{OTE}^l \times \overline{AE}^l, \quad (3.21)$$

where

$$\overline{OTE}^l \equiv \sum_{k=1}^{n_l} OTE^{l,k}(x^{l,k}, y^{l,k}) \times S^{l,k}, \quad (3.22)$$

and

$$\overline{OAE}^l \equiv \sum_{k=1}^{n_l} OAE^{l,k}(x^{l,k}, y^{l,k}, p) \times S_{ae}^{l,k}, \quad (3.23)$$

where

$$S^{l,k} \equiv \frac{py^{l,k}}{p\bar{Y}^l}, \quad S_{ae}^{l,k} \equiv \frac{p(y^{l,k}OTE^{l,k}(x^{l,k}, y^{l,k}))}{p \sum_{k=1}^{n_l} (y^{l,k}OTE^{l,k}(x^{l,k}, y^{l,k}))}, \quad k = 1, \dots, n_l. \quad (3.24)$$

In the next sub-section we provide some intuition behind these important results from which many other results can be derived.

3.4 Understanding the Fundamental Aggregation Results

Before going further, it is worth making a few intuitive remarks that should help in clarifying the fundamental results on aggregation summarized in the previous sub-section.

First of all, it should be clear that if $L = 1$ then the aggregate measures above are the efficiency measures for the entire group.

Second, note that the measure (3.22) can be viewed as a *multi-output* generalization of Farrell's measure of "*Structural Efficiency of an Industry*," (Farrell (1957), p. 261-262).

Third, recall that in the context of aggregation over industries, Domar (1961) derived a similar weighting scheme, using different arguments than outlined here and after imposing more restrictive assumptions. Meanwhile, Li and Ng (1995) proposed the same weights and decomposition of the aggregate revenue (although defined differently than above) into aggregate technical efficiency and aggregate allocative efficiency measures.¹³

Fourth, note that while the *technical* efficiency is constructed to be a price independent measure of efficiency, the aggregation weights for obtaining the (sub)group *technical* effi-

¹³To be precise, Li and Ng (1995) used a similar framework, yet without explicit relationship to the maximal revenue defined on the sum of the output sets and without noticing the theoretical link via the analogue of Koopmans (1957) theorem, and focusing on the DEA framework.

ciency derived above *depend* on prices. This might be viewed as undesirable. On the other hand, note that these weights were *not* chosen arbitrarily or in an *ad hoc* way, but came out as a result of imposing an economic criterion of optimizing behavior, which researchers also often consider as a benchmark when making their choice of orientation in measuring efficiency. Intuitively, if one wants to account for an *economic* importance of a DMU that obtained the particular ‘standardized’ efficiency score then, since prices contain important economic information, it shall not be surprising that the weights derived using the *economic* optimization principle are price-dependent. Another consideration is more practical: price information may be unavailable (or unreliable) in a given study. To circumvent this problem, one may use the shadow prices (Li and Ng (1995)). Alternatively, one may impose an extra assumption to make the derived weights price-independent, as we outline in sub-section 3.6.

Fifth, a condition often referred to as the ‘Law of One Price’ was assumed to enable feasibility of the derivations of these aggregation results. Importantly, note that this is a *necessary* assumption for obtaining a positive result in the stated aggregation problem. To be more precise, it is a necessary condition to establish an equivalence between the aggregate notions of efficiency (defined with respect to the aggregate technology and optimized as a group) and the dis-aggregate notions of efficiency (defined with respect to the individual technologies and optimized independently by each individual in the group). In other words, this ‘Law of One Price’ condition can be viewed as the condition of an equilibrium that ensures the system reaches the same outcome whether optimized individually and then aggregated or optimized over the aggregate technology by a group (e.g., a ‘central planner’ for the group). In this sense, the weights derived from this framework can be viewed as ‘*optimal* weights’, in the sense that they are derived from a framework where the system has reached equivalent optimal outcomes from both the aggregate and the dis-aggregate sides.

On the other hand, without this assumption, the general impossibility theorems of Blackorby and Russell (1999) are in action, which lead to much more disappointing conclusions for practitioners (since they ensure the impossibility of the equivalence) than this quite common condition in economic theory. Indeed, this condition is coherent with many economic models (perfect competition, Cournot-type oligopoly, etc.), where the notion of economic equilibrium indeed implies a common price. As many other theoretical assumptions it is, of course, simplifying the reality (e.g., see Kuosmanen et al. (2006) and Färe and Karagiannis (2017) for a discussion). In practice, it is of course possible to use the same formulas for the weights but with different prices and then compute an aggregate by averaging the individual efficiencies using such ‘non-optimal’ (or *ad hoc*) weights, and they can be viewed as approximations of the ‘optimal weights’ derived above. The problem is that such an aggregate is not guaranteed to be equivalent to the aggregate obtained with respect to the aggregate

technology, yet it may have another useful meaning that might be appealing from another perspective (e.g., it can be regarded as an aggregate efficiency that accounts for the price variation across the observations and thus showing the gap relative to the aggregate based on the ‘optimal weights’).

Finally, it should not be surprising that establishing *positive* aggregation results in economics requires extra and perhaps relatively strict assumptions. A good example would be the fairly strong conditions imposed to obtain the well-known in economic theory solutions to aggregation of demands, whether over goods or over consumers. Similarly, and as mentioned above, in the context of efficiency analysis, Blackorby and Russell (1999) analyzed a more general aggregation problem (without considering optimization behavior) and arrived at several *impossibility* results, concluding that very strong assumptions on the technology are needed for establishing positive aggregation results. The approach summarized above circumvents such assumptions by resorting to the optimization behavior (as a benchmark against which inefficiency of actual performance is measured) along with the other assumptions described above.

3.5 Aggregation of Aggregates

We now look at the case when a researcher wants to aggregate further, over already aggregate efficiency scores, i.e., across some sub-groups *within* a larger group. For example, suppose there is some partitioning of interest of the entire group into L non-intersecting and exhaustive sub-groups $l = 1, \dots, L$. Let $\bar{Y} \equiv \sum_{k=1}^n y^k = \sum_{l=1}^L \sum_{k=1}^{n_l} y^{l,k}$ be the total output across all DMUs in all the sub-groups. Also let the input allocation among firms *within* all the groups be denoted with $X = (X^1, \dots, X^L)$. If (3.11) is true for all $l = 1, \dots, L$, then we must have

$$\bar{P}(X) = \sum_{\oplus k=1}^n P^k(x^k) = \sum_{\oplus l=1}^L P^l(X^l) = \sum_{\oplus l=1}^L \sum_{\oplus k=1}^{n_l} P^{l,k}(x^{l,k}), \quad (3.25)$$

i.e., the aggregate output set of all groups together is the Minkowski sum of the group output sets, over $l = 1, \dots, L$. Thus, $\bar{P}(X)$ would inherit its properties from the properties of sub-group technologies, which in turn are inherited from the regularity conditions imposed on the output sets of individual DMUs.

Using the group technology (3.25), one can define the *group revenue* function as

$$\bar{R}(X, p) \equiv \sup_y \{py : y \in \bar{P}(X)\}, \quad (3.26)$$

which, similarly to (3.7), gives rise to the *group revenue efficiency* measure

$$\overline{RE}(X, \bar{Y}, p) \equiv \bar{R}(X, p) / p\bar{Y}. \quad (3.27)$$

An immediate consequence of the previous theorem and of (3.25) is summarized in the next corollary.

Corollary 3. *The maximal revenue of the entire group of DMUs feasible from (X, p) is equal to the sum of maximal revenues of all its (non-intersecting) sub-groups of DMUs feasible from (X^l, p) , $l = 1, \dots, L$, i.e.,*

$$\bar{R}(X, p) = \sum_{l=1}^L \bar{R}^l(X^l, p) = \sum_{l=1}^L \sum_{k=1}^{n_l} R^{l,k}(x^{l,k}, p). \quad (3.28)$$

The intuition of this result is the same as that of its analogue of (3.18)—it is its extension to the aggregation between the sub-groups into a larger group. The corresponding result about aggregation of revenue efficiency measures is summarized in the next corollary.

Corollary 4. *We have*

$$\overline{RE}(X, \bar{Y}, p) = \sum_{l=1}^L \overline{RE}^l(X^l, \bar{Y}^l, p) \times \mathcal{S}^l, \quad (3.29)$$

where

$$\mathcal{S}^l = p\bar{Y}^l / \left(p \sum_{l=1}^L \bar{Y}^l \right), \quad l = 1, \dots, L. \quad (3.30)$$

Intuitively, this corollary says that the weighted sum of the revenue efficiencies of revenue-maximizing sub-groups of firms is the same as the revenue efficiency with respect to the aggregate technology (3.25) for the group that unites these sub-groups (assuming all firms face the same output prices and the standard regularity conditions hold). That is, it is an analogue of (3.19)-(3.20). In turn, this corollary implies the following important result.

Corollary 5. *We have*

$$\overline{RE}(X, \bar{Y}, p) = \overline{OTE} \times \overline{OAE}, \quad (3.31)$$

where

$$\overline{OTE} = \sum_{l=1}^L \overline{OTE}^l \times \mathcal{S}^l, \quad (3.32)$$

and

$$\overline{OAE} = \sum_{l=1}^L \overline{OAE}^l \times \mathcal{S}_{ae}^l, \quad (3.33)$$

where

$$\mathcal{S}^l = p\bar{Y}^l / \left(p \sum_{l=1}^L \bar{Y}^l \right), \quad l = 1, \dots, L, \quad (3.34)$$

and

$$\mathcal{S}_{ae}^l = \left(p\bar{Y}^l \times \overline{OTE}^l \right) / \left(p \sum_{l=1}^L \bar{Y}^l \times \overline{OTE}^l \right), \quad l = 1, \dots, L. \quad (3.35)$$

Intuitively, this last corollary provides a theoretically justified weighting scheme for an aggregation over sub-groups of the aggregate technical and aggregate allocative efficiencies into more aggregate analogues, and such that they decompose the aggregated revenue efficiency. Thus, this approach provides ‘internally consistent’ aggregation *within* and *between* the sub-groups.

3.6 Price Independent Weights

In this section we summarize the method for converting the derived above price-dependent weights into the price-independent weight such that the same aggregation scheme based on and derived from the economic principles is preserved. This method was proposed by Färe and Zelenyuk (2003, 2007) and Simar and Zelenyuk (2007). We first focus on the case of aggregating efficiency scores of the *entire* group. The key additional assumption here is the following:

$$p_m \bar{Y}_m / \left(\sum_{m=1}^M p_m \bar{Y}_m \right) = \alpha_m, \quad m = 1, \dots, M, \quad (3.36)$$

where $\bar{Y}_m \equiv \sum_{k=1}^n y_m^k$ and $\alpha_m \in (0, 1)$ is a constant (known or estimated) for all $m \in \{1, \dots, M\}$ with normalization $\sum_{m=1}^M \alpha_m = 1$. Intuitively, (3.36) states that the weight of the *industry* revenue from the output m in the industry total revenue equals α_m . Furthermore, let us denote $\varpi_m^k = y_m^k / \bar{Y}_m$ to be the weight of the k^{th} firm in the group in terms of the m^{th} -output, and let us impose the condition (3.36) upon the weights for the aggregation of the revenue and technical efficiency scores derived above, to obtain

$$S^k = \sum_{m=1}^M \alpha_m \varpi_m^k, \quad k = 1, \dots, n. \quad (3.37)$$

Intuitively, (3.37) says that the weight of a firm is the weighted average over all the output-shares of this firm in its group, where the weights are the revenue shares of the industry for each output m in the total revenue of the industry. Next, use (3.36) and (3.34) to derive the

weights for aggregating ‘*between* the sub-groups’

$$\mathcal{S}^l = \sum_{m=1}^M \alpha_m W_m^l, \quad l = 1, \dots, L, \quad (3.38)$$

where $W_m^l = \bar{Y}_m^l / \bar{Y}_m$ is the share of the l^{th} sub-group in the entire group in terms of the m^{th} -output. Furthermore, with a bit more algebra we can derive the *price-independent weight* for an individual efficiency of firm k ‘*within* a sub-group l ’ to be

$$S^{l,k} = S^k / \mathcal{S}^l, \quad k = 1, \dots, n_l; \quad l = 1, \dots, L, \quad (3.39)$$

i.e., we get an analogue of (3.37) which accounts for the weight of each particular sub-group in the entire group.

On the other hand, the price-independent weights for aggregating *allocative* efficiencies are derived similarly as above but where the observed outputs are replaced with their *technically-efficient* analogues, i.e.,

$$S_{ae}^k = \frac{S^k \times OTE^k(x^k, y^k)}{\sum_{k=1}^n S^k \times OTE^k(x^k, y^k)}, \quad k = 1, \dots, n, \quad (3.40)$$

where S^k is given in (3.37).

Meanwhile, we can also employ the standardization in (3.36), along with (3.34), to get the weights for aggregating ‘*between* the sub-groups’ to be given by

$$\mathcal{S}_{ae}^l = \frac{\sum_{m=1}^M \alpha_m W_m^l \overline{OTE}^l}{\sum_{l=1}^L \sum_{m=1}^M \alpha_m W_m^l \overline{OTE}^l} = \frac{\overline{OTE}^l \times \mathcal{S}^l}{\sum_{l=1}^L \overline{OTE}^l \times \mathcal{S}^l}, \quad l = 1, \dots, L, \quad (3.41)$$

where $W_m^l = \bar{Y}_m^l / \bar{Y}_m$ is the weight of l^{th} sub-group in the entire group in terms of the m^{th} -output. Note that (3.41) is analogous to what we obtained for the individual firms, but for the sub-group level. Moreover, (3.41) can be used to derive the weight of an individual efficiency of firm k ‘*within* a sub-group l ’ to be

$$S_{ae}^{l,k} = S_{ae}^k / \mathcal{S}_{ae}^l, \quad k = 1, \dots, n_l; \quad l = 1, \dots, L.$$

Finally, note that all the derivations here were done for the case of the output orientation and analogous derivations can be made for the case of input orientation (and, potentially, for the joint input-output or profit-orientation), which we leave as exercises for the readers (see Färe et al. (2004b) and Mayer and Zelenyuk (2014a,b) for some related derivations).

4 Aggregation of Productivity Indexes

Similarly as with the efficiency scores, applied studies involving productivity indexes usually need to present some aggregates of the estimated productivity indexes—to summarize the overall tendencies in a sample, to perform statistical inference about the population, etc. Typically, researchers use the *simple* or the *equally-weighted geometric* mean for this purpose. The discussion above suggests that it would also be important to have *some well-justified weights* when aggregating *productivity indexes*. Such weights would help in accounting for the relative importance of each firm whose index is entering into the average. This question was first addressed by Zelenyuk (2006), who derived an aggregation scheme for the Malmquist productivity index (MPI), and we summarize this approach in this section.¹⁴ To simplify the notation, from now on we will consider just one group, i.e., drop the sub-group subscript l (but add the time subscript $\tau = s, t$).

4.1 Individual Malmquist Productivity Indexes

Let us first recall the definitions of the MPI. We will focus on measuring changes in productivity from a period s to a period t ($s < t$). Recall that the output oriented MPI can be defined as

$$M^k(y_s^k, y_t^k, x_s^k, x_t^k) \equiv \left[\frac{D_s^k(x_t^k, y_t^k)}{D_s^k(x_s^k, y_s^k)} \times \frac{D_t^k(x_t^k, y_t^k)}{D_t^k(x_s^k, y_s^k)} \right]^{1/2}. \quad (4.1)$$

where $D_s^k(x_t^k, y_t^k)$ is the Shephard's output oriented distance function that we now cast in the inter-temporal framework, characterizing technology of DMU k in period s and evaluated at the point (x_t^k, y_t^k) . Note that we dropped the subscript “o” to simplify our already intense notation.¹⁵

In the light of the duality between the distance function and the revenue function, one can also define the *revenue* (or dual) analogue of the *MPI* as

$$\begin{aligned} RM^k(\cdot) &\equiv RM^k(p_s, p_t, y_s^k, y_t^k, x_s^k, x_t^k) \\ &\equiv \left[\left(\frac{RE_s^k(x_t^k, y_t^k, p_t)}{RE_s^k(x_s^k, y_s^k, p_s)} \times \frac{RE_t^k(x_t^k, y_t^k, p_t)}{RE_t^k(x_s^k, y_s^k, p_s)} \right)^{-1} \right]^{1/2}, \end{aligned} \quad (4.2)$$

which, naturally, can be decomposed as

$$RM^k(\cdot) \equiv M^k(\cdot) \times AM^k(\cdot), \quad (4.3)$$

¹⁴Also see Mayer and Zelenyuk (2014a, 2017) for extensions of this approach.

¹⁵Again, here we focus on the output orientation case and similar developments can be done for the input orientation case. See Mayer and Zelenyuk (2014a, 2017) for some of these details.

where $M^k(\cdot)$ is defined in (4.1), and $AM^k(\cdot)$ is the allocative component of the dual MPI, defined as

$$\begin{aligned} AM^k(\cdot) &\equiv AM^k(p_s, p_t, y_s^k, y_t^k, x_s^k, x_t^k) \\ &\equiv \left[\left(\frac{OAE_s^k(x_t^k, y_t^k, p_t)}{OAE_s^k(x_s^k, y_s^k, p_s)} \times \frac{OAE_t^k(x_t^k, y_t^k, p_t)}{OAE_t^k(x_s^k, y_s^k, p_s)} \right)^{-1} \right]^{1/2}. \end{aligned} \quad (4.4)$$

4.2 Aggregation Problem: Inter-temporal Perspective

Here we adapt the aggregation concepts outlined above to the inter-temporal framework. As for the case of efficiency aggregation, a key stepping-stone for deriving the aggregation results for productivity indexes is to define a relevant *group technology*, and as before, here we admit the *additive structure of aggregation of the output sets*, i.e.,

$$\bar{P}_\tau(X) \equiv \sum_{\oplus k=1}^n P_\tau^k(x^k), \quad \tau = s, t, \quad (4.5)$$

and so the *group revenue* function at period τ is given by

$$\bar{R}_\tau(X, p) \equiv \max_y \{py \quad : \quad y \in \bar{P}_\tau(X)\}, \quad \tau = s, t, \quad (4.6)$$

while the revenue efficiency at τ is given by

$$\overline{RE}_\tau(X, \bar{Y}, p) \equiv \bar{R}_\tau(X, p) / p\bar{Y}, \quad \tau = s, t. \quad (4.7)$$

Now, to measure changes in productivity between s and t , let the *group* (aggregate) analog of (4.2) be

$$\overline{RM}(p_s, p_t, \bar{Y}_s, \bar{Y}_t, X_s, X_t) \equiv \left[\left(\frac{\overline{RE}_s(X_t, \bar{Y}_t, p_t)}{\overline{RE}_s(X_s, \bar{Y}_s, p_s)} \times \frac{\overline{RE}_t(X_t, \bar{Y}_t, p_t)}{\overline{RE}_t(X_s, \bar{Y}_s, p_s)} \right)^{-1} \right]^{\frac{1}{2}}, \quad (4.8)$$

where the time subscripts indicate the particular values of efficiency measures for specific periods $\tau = s, t$.

Ideally, we want to find an aggregation function $f_{RM}(\cdot)$ that can relate the aggregate measure (4.8) to the individual measures (4.2) in some ‘meaningful’ way. Being unable to find such a ‘meaningful’ way, Zelenyuk (2006) resorted to something that may seem ‘less than ideal’, yet feasible—find an aggregation function $f_{RE}(\cdot)$ that can relate the aggregate measure (4.8) to all the components of all the individual measures (4.2), in a ‘meaningful’

way, i.e., so that we have

$$\overline{RM}(p_s, p_t, \overline{Y}_s, \overline{Y}_t, X_s, X_t) = f_{RE}(RE_\tau^1(\cdot), \dots, RE_\tau^n(\cdot)), \quad \tau = s, t, \quad (4.9)$$

such that, preferably, the decomposition (4.3) is maintained at the *aggregate* level, i.e.,

$$\overline{RM}(p_s, p_t, \overline{Y}_s, \overline{Y}_t, X_s, X_t) = \overline{M}(\cdot) \times \overline{AM}(\cdot), \quad (4.10)$$

where, in turn, one need to find some aggregation functions $f_D(\cdot)$, $f_{AE}(\cdot)$ that ensure that the aggregate *primal* MPI is related to all the components of all the individual analogues (4.1), i.e.,

$$\overline{M}(\cdot) \equiv \overline{M}(\overline{Y}_s, \overline{Y}_t, X_s, X_t) \equiv f_D(D_\tau^1(\cdot), \dots, D_\tau^n(\cdot)), \quad \tau = s, t, \quad (4.11)$$

while the aggregate *allocative*-MPI is related to (4.4) or its individual components, i.e.,

$$\overline{AM}(\cdot) \equiv \overline{AM}(\overline{Y}_s, \overline{Y}_t, X_s, X_t) \equiv f_A(OAE_\tau^1(\cdot), \dots, OAE_\tau^n(\cdot)), \quad \tau = s, t. \quad (4.12)$$

Such functions are found in the next sub-section using, again, the Koopmans-type arguments that we described above.

4.3 Aggregation of the MPIs

As for the aggregation of efficiency scores, the foundation here is the inter-temporal extension of the aggregation theory from Färe and Zelenyuk (2003), which in turn is an adaptation of aggregation theory from Koopmans (1957), which we restate below casting it in the inter-temporal framework, with the time subscript $\tau = s, t$:

$$\overline{R}_\tau(X, p) = \sum_{k=1}^n R_\tau^k(x^k, p), \quad x^k \in \mathfrak{R}_+^N, \quad \forall k = 1, \dots, n, \quad p \in \mathfrak{R}_{++}^M, \quad (4.13)$$

and therefore, for $j, \tau = s, t$, we have

$$\overline{RE}_\tau(X_j, \overline{Y}_j, p_j) = \sum_{k=1}^n RE_\tau^k(x_j^k, y_j^k, p_j) \times S_j^k \quad (4.14)$$

where

$$S_j^k \equiv p_j y_j^k / p_j \overline{Y}_j, \quad k = 1, \dots, n. \quad (4.15)$$

Thus, the decomposition is maintained at the aggregate level: for any $j, \tau = s, t$, we have

$$\overline{RE}_\tau(X_j, \overline{Y}_j, p_j) = \overline{OTE}_\tau(j) \times \overline{OAE}_\tau(j), \quad (4.16)$$

where

$$\overline{OTE}_\tau(j) \equiv \sum_{k=1}^n [D_\tau^k(x_j^k, y_j^k)]^{-1} \times S_j^k, \quad (4.17)$$

$$\overline{OAE}_\tau(j) \equiv \sum_{k=1}^n OAE_\tau^k(x_j^k, y_j^k, p_j) \times S_{ae,\tau,j}^k, \quad (4.18)$$

and

$$S_{ae,\tau,j}^k \equiv \frac{p_j (y_j^k / D_\tau^k(x_j^k, y_j^k))}{p_j \sum_{k=1}^n (y_j^k / D_\tau^k(x_j^k, y_j^k))}, \quad k = 1, \dots, n. \quad (4.19)$$

Furthermore, applying (4.8), (4.14) and (4.15) we get a desired aggregation result—a solution to (4.9), given by:

$$\begin{aligned} & \overline{RM}(p_s, p_t, \bar{Y}_s, \bar{Y}_t, X_s, X_t) \\ &= \left[\left(\frac{\sum_{k=1}^n RE_s^k(x_t^k, y_t^k, p_t) \times S_t^k}{\sum_{k=1}^n RE_s^k(x_s^k, y_s^k, p_s) \times S_s^k} \right. \right. \\ & \quad \left. \left. \times \frac{\sum_{k=1}^n RE_t^k(x_t^k, y_t^k, p_t) \times S_t^k}{\sum_{k=1}^n RE_t^k(x_s^k, y_s^k, p_s) \times S_s^k} \right)^{-1} \right]^{1/2}. \end{aligned} \quad (4.20)$$

Importantly, note that the decomposition at the *aggregate* level is preserved and given by

$$\overline{RM}(p_s, p_t, \bar{Y}_s, \bar{Y}_t, X_s, X_t) = \overline{M}(\cdot) \times \overline{AM}(\cdot), \quad (4.21)$$

where the solutions to (4.11) and (4.12) are given, respectively, by

$$\overline{M}(\bar{Y}_s, \bar{Y}_t, X_s, X_t) = \left[\left(\frac{\overline{OTE}_s(t)}{\overline{OTE}_s(s)} \times \frac{\overline{OTE}_t(t)}{\overline{OTE}_t(s)} \right)^{-1} \right]^{1/2}, \quad (4.22)$$

and

$$\overline{AM}(p_s, p_t, \bar{Y}_s, \bar{Y}_t, X_s, X_t) = \left[\left(\frac{\overline{OAE}_s(t)}{\overline{OAE}_s(s)} \times \frac{\overline{OAE}_t(t)}{\overline{OAE}_t(s)} \right)^{-1} \right]^{1/2}, \quad (4.23)$$

where, in turn, the four components inside (4.22) are given in (4.17) while the four components inside (4.23) are given in (4.18).

The theoretical and practical importance of these results is that they give explicit formulas for aggregation of the MPIs. In particular, they give a way of obtaining a *group* productivity change score from the *individual* analogues, where the aggregation function and the aggregation weights are not *ad hoc* but derived from economic principles, besides being intuitive.

4.4 Geometric vs. Harmonic Averaging

In earlier studies, noting on the multiplicative nature of the MPI, researchers often used not only the equal weights but also the *geometric* rather than the arithmetic or the harmonic averaging of the *individual* estimates when they wished to summarize the point-estimates of MPIs (e.g., see Färe et al. (1994)). That is, not only the weights were equal, but also the aggregating function used in previous practice was quite different from what the theoretical derivations in the previous sub-section suggested. How can these different approaches be reconciled? From the discussions above, it must be clear that the weights can dramatically influence the results, whether quantitatively or qualitatively. A natural question is whether the functional form of the aggregation is critical and, in particular, can one use the geometric mean rather than the arithmetic mean?

This question was also addressed by Zelenyuk (2006), who pointed out that (4.22) can be restated in terms of *harmonic* aggregations of individual distance functions, i.e.,

$$\begin{aligned} \overline{M}(\cdot) = & \left[\frac{(\sum_{k=1}^n [D_s^k(x_t^k, y_t^k)]^{-1} \times S_t^k)^{-1}}{(\sum_{k=1}^n [D_s^k(x_s^k, y_s^k)]^{-1} \times S_s^k)^{-1}} \right. \\ & \left. \times \frac{(\sum_{k=1}^n [D_t^k(x_t^k, y_t^k)]^{-1} \times S_t^k)^{-1}}{(\sum_{k=1}^n [D_t^k(x_s^k, y_s^k)]^{-1} \times S_s^k)^{-1}} \right]^{\frac{1}{2}}. \end{aligned} \quad (4.24)$$

and its geometric analogue can be defined as

$$\overline{M}^G(\cdot) \equiv \left[\frac{\prod_{k=1}^n D_s^k(x_t^k, y_t^k)^{\omega_t^k}}{\prod_{k=1}^n D_s^k(x_s^k, y_s^k)^{\omega_s^k}} \times \frac{\prod_{k=1}^n D_t^k(x_t^k, y_t^k)^{\omega_t^k}}{\prod_{k=1}^n D_t^k(x_s^k, y_s^k)^{\omega_s^k}} \right]^{1/2}, \quad (4.25)$$

for some weights ω_t^k, ω_s^k .

The aggregation usually used in practice is a particular case of (4.25) that assumes equal weights across all k , i.e., $\omega_t^k = \omega_s^k = 1/n$. It must be clear that, in general, (4.24) is not equal to (4.25) and, in fact, no *exact* general relationship exists between the two. However, taking the first-order approximation of $\prod_{k=1}^n D_s^k(x_t^k, y_t^k)^{S_t^k}$ and of $(\sum_{k=1}^n [D_s^k(x_t^k, y_t^k)]^{-1} \times S_t^k)^{-1}$ around *unity* (which is a natural point around which productivity and efficiency indexes can be approximated) in both cases we get $\sum_{k=1}^n D_s^k(x_t^k, y_t^k) \times S_t^k$, meaning that one can conclude

$$\overline{M}(\overline{Y}_s, \overline{Y}_t, X_s, X_t) \cong \overline{M}^G(\overline{Y}_s, \overline{Y}_t, X_s, X_t), \text{ for } (\omega_t^k, \omega_s^k) = (S_t^k, S_s^k). \quad (4.26)$$

In words, (4.26) states that the *first-order-approximation* relationship exists between the

aggregate MPI constructed with harmonic components derived above and the *geometric* aggregate of individual MPI's, if both use the same set of weights. This implies that, for anyone who prefers the geometric aggregation, this relationship gives a justification for choosing the aggregation weights (which are more influential)—the weights derived from economic principles, which account for the economic weight of each firm.

A natural question is “How substantial is the difference between the geometric and harmonic aggregations?”¹⁶ Zelenyuk (2006) presented some simulation results confirming that the difference is fairly small. For instance, if the scores of a productivity index come from uniform distribution around unity with the range of 50 percentage points (thus allowing a substantial change), then the square-root of the mean squared difference between the harmonic and geometric means across various simulations was only about 1 percentage point. Thus, a practical implication that one can deduce from here is that the geometric-type and the harmonic-type aggregations of the productivity indexes (under the same weights) give similar aggregate scores for moderate variations of the scores being aggregated. In other words, the aggregation function per se (whether geometric, harmonic or arithmetic) is not as crucial—what is more important are the weights of aggregation, which needs to be justified on some theoretical grounds.¹⁷

4.5 Decomposition and Aggregation

The aggregation results we summarized above can also be extended to the aggregation of components of various decompositions of MPIs. While there are many decompositions of MPI offered in the literature, here we focus on what seems to be the most popular decomposition in practice—the one proposed in the seminal work of Färe et al. (1994), as the following

$$M^k(\cdot) \equiv EFCH^k(\cdot) \times TECH^k(\cdot), \quad (4.27)$$

¹⁶From theory, it is known that under the same weighting scheme, the geometric mean is larger than the harmonic mean but smaller than the arithmetic mean. Note however that the aggregate MPI in (4.24) involves products of ratios of the harmonic means and so it can be smaller or greater than the aggregate MPI obtained via a geometric mean as in (4.25), depending on the relative magnitudes that appear in the numerators and denominators of (4.24). Both means are approximately equal (to the arithmetic mean) in the sense of first order approximation around unity.

¹⁷One should however be careful aggregating when there are scores equal or very close to zero: both geometric and harmonic averages completely fail if at least one element is zero and may yield an unreasonably low aggregate score if at least one element is very close to zero (even if many others have large efficiency or productivity scores), unless they are ‘neutralized’ by a very low weight in the aggregation, as can be done with weighted aggregates. In such cases, using arithmetic aggregation, which is less sensitive to the outliers, could also be a better solution.

where the first component is referred to as the *efficiency change*, defined as

$$EFCH^k(\cdot) \equiv EFCH^k(y_s^k, y_t^k, x_s^k, x_t^k) \equiv \frac{D_t^k(x_t^k, y_t^k)}{D_s^k(x_s^k, y_s^k)}, \quad (4.28)$$

and the second component is referred to as the *technological change*, defined as

$$TECH^k(\cdot) \equiv TECH^k(y_s^k, y_t^k, x_s^k, x_t^k) \equiv \left[\frac{D_s^k(x_t^k, y_t^k)}{D_t^k(x_t^k, y_t^k)} \times \frac{D_s^k(x_s^k, y_s^k)}{D_t^k(x_s^k, y_s^k)} \right]^{1/2}. \quad (4.29)$$

The aggregation question then is to find appropriate *group* analogues to (4.28) and (4.29), i.e., some functions $f_{EC}(\cdot)$ and $f_{TC}(\cdot)$ that relate the aggregate measures to individual ones. As above, a natural choice is to utilize the Koopmans' type arguments, as was done in Zelenyuk (2006), to arrive at

$$\overline{EFCH}(\cdot) = \left(\frac{\overline{OTE}_t(t)}{\overline{OTE}_s(s)} \right)^{-1} = \frac{(\sum_{k=1}^n [D_t^k(x_t^k, y_t^k)]^{-1} \times S_t^k)^{-1}}{(\sum_{k=1}^n [D_s^k(x_s^k, y_s^k)]^{-1} \times S_s^k)^{-1}}, \quad (4.30)$$

and

$$\begin{aligned} \overline{TECH}(\cdot) &= \left[\left(\frac{\overline{OTE}_s(t)}{\overline{OTE}_t(t)} \times \frac{\overline{OTE}_s(s)}{\overline{OTE}_t(s)} \right)^{-1} \right]^{1/2} \\ &= \left[\frac{(\sum_{k=1}^n [D_s^k(x_t^k, y_t^k)]^{-1} \times S_t^k)^{-1} (\sum_{k=1}^n [D_s^k(x_s^k, y_s^k)]^{-1} \times S_s^k)^{-1}}{(\sum_{k=1}^n [D_t^k(x_t^k, y_t^k)]^{-1} \times S_t^k)^{-1} (\sum_{k=1}^n [D_t^k(x_s^k, y_s^k)]^{-1} \times S_s^k)^{-1}} \right]^{1/2}. \end{aligned} \quad (4.31)$$

As before, the first order approximation relationship can also be established between the harmonic-type aggregations in (4.30) and (4.31) and their geometric analogues. Moreover, these aggregation results can also be extended to aggregation across or over larger groups, in a similar manner as for aggregating efficiency scores that we discussed above, i.e., extending Simar and Zelenyuk (2007).

5 Aggregation for Scale Measures

Measurement of economies of scale for an individual firm or for an industry has been one of the most frequently addressed research questions in economics and applied econometrics. This is usually done via estimating such measures as scale elasticity and/or scale efficiency. Here we will focus on the elasticity approach, following Färe and Zelenyuk (2012), while the

aggregation for scale efficiency can be found in Zelenyuk (2015).

For analyzing economies of scale for a group (e.g., industry or sub-industry), researchers usually estimate the elasticity *at* some points of interest, e.g., the non-weighted mean or the median of the data or, alternatively, the non-weighted mean of the *individual* estimates of scale elasticities. Importantly, note that these different approaches do not give the same information, in general, and each have certain theoretical or practical appeals and *caveats*. Here we discuss another theoretical approach of measuring scale elasticity of a *group*, which is based on a similar aggregation result as that derived above.

So far we considered the output oriented framework and in this section, because researchers often focus on elasticity of the cost function, we will consider the case of input orientation. To do so, first note that the technology set of firm k can be equivalently characterized by the input requirement sets, defined as

$$\mathcal{L}^k(y^k) \equiv \{x : x \text{ can produce } y^k\}, \quad y \in \mathfrak{R}_+^M, \quad (5.1)$$

so that $x^k \in \mathcal{L}^k(y^k)$, $y \in \mathfrak{R}_+^M \iff (x^k, y^k) \in T^k$. In turn, technology can also be equivalently characterized by the *input oriented* Shephard (1953) distance function $D_i^k : \mathfrak{R}_+^M \times \mathfrak{R}_+^N \rightarrow \mathfrak{R}_+ \cup \{\infty\}$, defined as

$$D_i^k(y^k, x^k) \equiv \sup_{\delta} \{\delta > 0 : x^k / \delta \in \mathcal{L}^k(y^k)\}. \quad (5.2)$$

A closely related concept is the input oriented *Farrell* measure of *technical efficiency*,

$$ITE^k(y^k, x^k) \equiv \inf_{\theta} \{\theta > 0 : \theta x^k \in \mathcal{L}^k(y^k)\} = 1/D_i^k(y^k, x^k). \quad (5.3)$$

In addition to assuming the main regularity axioms of production theory (A1-A5), we also assume convexity of the input requirement sets, i.e.,

A7: Input requirement sets $\mathcal{L}^k(y^k)$ are convex, $\forall y^k \in \mathfrak{R}_+^M$.

As a result, due to duality theory in economics (see Shephard (1953); Färe and Primont (1995) and Sickles and Zelenyuk (2019)), the technology can be equivalently characterized by the *cost* function, $C^k : \mathfrak{R}_+^M \times \mathfrak{R}_{++}^N \rightarrow \mathfrak{R}_+ \cup \{\infty\}$, defined as

$$C^k(y^k, w) \equiv \inf_x \{wx : x \in \mathcal{L}^k(y^k)\}, \quad (5.4)$$

where $w \equiv (w_1, \dots, w_N) \in \mathfrak{R}_{++}^N$ is the vector of input prices. The related efficiency measure

for a DMU k in the *dual* input-oriented framework would then be the *cost efficiency* (also referred to as or the overall input efficiency), defined formally as,

$$CE^k(y^k, x^k, w) \equiv C^k(y^k, w)/wx^k. \quad (5.5)$$

From the duality theory for the cost function, we also have the so-called Mahler's inequality,

$$C^k(y^k, w) \geq wx^k/D_i^k(y^k, x^k), \quad (5.6)$$

leading to the notion of the input oriented *allocative* (in)efficiency, defined as a multiplicative residual that turns (5.6) into equality, i.e.,

$$IAE^k(y^k, x^k, w) \equiv CE^k(y^k, x^k, w)/ITE^k(y^k, x^k), \quad (5.7)$$

and so we have another useful decomposition:

$$CE^k(y^k, x^k, w) = ITE^k(y^k, x^k) \times IAE^k(y^k, x^k, w), \quad (5.8)$$

We will use these efficiency measures and their aggregate analogues later in the chapter, while for the framework of scale elasticity we will focus on the distance function and the cost function characterizations.

It is important to note that one can use both the primal and the dual characterizations to measure *economies of scale* via the *scale elasticity*. Specifically, with appropriate differentiability assumptions, for the *dual* framework the scale elasticity is defined as¹⁸

$$e_c(y^k, w) \equiv \left. \frac{\partial \ln C^k(y^k \theta, w)}{\partial \ln \theta} \right|_{\theta=1} = \frac{\nabla'_{y^k} C^k(y^k, w) y^k}{C^k(y^k, w)}, \quad (5.9)$$

and for the *primal* framework, the scale elasticity is defined as¹⁹

$$e_i(y^k, x^k) \equiv \left. \frac{\partial \ln \lambda}{\partial \ln \theta} \right|_{\substack{D_i^k(y^k \theta, x^k \lambda)=1, \\ \theta=1, \lambda=1}} = -\nabla'_{y^k} D_i^k(y^k, x^k) y^k. \quad (5.10)$$

Now, suppose x^{*k} is a solution to (5.4), then one can obtain equality between the dual and the primal measures, i.e.,²⁰

$$e_c(y^k, w) = e_i(y^k, x^{*k}), \quad (5.11)$$

¹⁸E.g., see Panzar and Willig (1977).

¹⁹See Färe et al. (1986); Färe and Primont (1995).

²⁰See Färe et al. (1986); Zelenyuk (2013b) for more details on this.

where x^{*k} is a solution to (5.4), i.e.,

$$x^{*k} \equiv \arg \inf_x \{wx \quad : \quad x \in \mathcal{L}^k(y^k)\}. \quad (5.12)$$

Intuitively, (5.11) states that the same information about the scale elasticity of an *individual* firm k can be obtained from the primal and dual approaches.

Now, analogous to what we did in previous sections, let the *group* input requirement set be given by the Minkowski sum of the individual input requirement sets across all DMUs k ($k = 1, \dots, n$), i.e.,

$$\overline{\mathcal{L}}(y^1, \dots, y^n) = \sum_{\oplus k=1}^n \mathcal{L}^k(y^k). \quad (5.13)$$

Note that $\overline{\mathcal{L}}(y^1, \dots, y^n)$ inherits the regularity conditions imposed on the individual input requirement sets and, in particular, convexity of the individual input requirement sets implies that $\overline{\mathcal{L}}(y^1, \dots, y^n)$ is also convex. Also note that the aggregation structure defined by $\overline{\mathcal{L}}(y^1, \dots, y^n)$ presumes no reallocation of outputs and no externalities across the individuals.

In turn, the *group* cost function would be the *aggregate* analogue of (5.4), defined as

$$\overline{C}(y^1, \dots, y^n, w) \equiv \inf_x \{wx \quad : \quad x \in \overline{\mathcal{L}}(y^1, \dots, y^n)\}, \quad (5.14)$$

while the *group* input oriented distance function would be defined as

$$\overline{D}_i(y^1, \dots, y^n, \sum_{k=1}^n x^k) \equiv \sup_{\delta} \{\delta > 0 \quad : \quad (\sum_{k=1}^n x^k / \delta) \in \overline{\mathcal{L}}(y^1, \dots, y^n)\}, \quad (5.15)$$

as the *aggregate* analogue of (5.2).

Therefore, one can measure the *economies of scale* for the *group* from the measures of scale elasticity defined for the aggregate technology— analogously to how it is done for the individual technologies, i.e., we have

$$\begin{aligned} \overline{e}_c(y^1, \dots, y^n, w) &\equiv \left. \frac{\partial \ln \overline{C}(y^1 \theta, \dots, y^n \theta, w)}{\partial \ln \theta} \right|_{\theta=1} \\ &= \frac{\nabla'_Y \overline{C}(y^1, \dots, y^n, w) Y}{\overline{C}(y^1, \dots, y^n, w)}, \end{aligned} \quad (5.16)$$

where $\nabla'_Y \overline{C}(y^1, \dots, y^n, w) \equiv (\partial \overline{C}(y^1, \dots, y^n, w) / \partial y^1, \dots, \partial \overline{C}(y^1, \dots, y^n, w) / \partial y^n)$ and $Y \equiv$

$(y^1, \dots, y^n)'$. Meanwhile, for the primal framework, we get

$$\begin{aligned} \bar{e}_i(y^1, \dots, y^n, \sum_{k=1}^n x^k) &\equiv \frac{\partial \ln \lambda}{\partial \ln \theta} \bigg|_{\substack{\bar{D}_i(y^1 \theta, \dots, y^n \theta, \sum_{k=1}^n x^k \lambda)=1, \\ \theta=1, \lambda=1}} \\ &= -\nabla'_Y \bar{D}_i(y^1, \dots, y^n, \sum_{k=1}^n x^k) Y. \end{aligned} \quad (5.17)$$

Furthermore, let x^* be a solution to (5.14), then the dual and the primal measures of *group* scale elasticity would be equal, i.e., putting this formally, we have a desired result:

$$\bar{e}_c(y^1, \dots, y^n, w) = \bar{e}_i(y^1, \dots, y^n, x^*), \quad (5.18)$$

where

$$x^* \equiv \arg \inf_x \{wx : x \in \bar{\mathcal{L}}(y^1, \dots, y^n)\}.$$

The reader shall notice that (5.18) is an aggregate analog of (5.11). The main goal therefore now is to find a relationship between the aggregate and the individual scale elasticity measures that will enable getting the aggregate measures from the individual ones. As above, the fundamental step for reaching this aim is the following result.

Theorem 2. *The minimal cost of the group of DMUs with production plan $Y = (y^1, \dots, y^n)$ is equal to the sum of the minimal costs of all its member DMUs with the same production plan y^1, \dots, y^n , assuming all the member DMUs face the same input prices w , i.e.,*

$$\bar{C}(y^1, \dots, y^n, w) = \sum_{k=1}^n C^k(y^k, w). \quad (5.19)$$

In words, this theorem states that whether the group of DMUs minimize the costs for their given output plans together via a ‘social planner’ (and without reallocation of outputs across DMUs) or they minimize individually and then these costs are summed over, the result should be the same if they face the same input prices w . This theorem is the cost analog of the theorem of Koopmans (1957) for aggregation of profit functions (see Färe et al. (2004b) for a proof).

Now, for measuring the change in costs due to infinitesimal and *equiproportional* change of *all* outputs, we differentiate both sides of (5.19) *along* the ray from the origin through the

point $Y \equiv (y^1, \dots, y^n)'$. Doing so for the l.h.s. of (5.19) we get

$$\left. \frac{\partial \bar{C}(y^1\theta, \dots, y^n\theta, w)}{\partial \theta} \right|_{\theta=1} = \nabla'_Y \bar{C}(y^1, \dots, y^n, w)Y. \quad (5.20)$$

while doing so for the r.h.s. of (5.19) we get

$$\left. \partial \left(\sum_{k=1}^n C^k(y^k\theta, w) \right) / \partial \theta \right|_{\theta=1} = \sum_{k=1}^n \nabla'_{y^k} C^k(y^k, w)y^k. \quad (5.21)$$

and combining the two, we get the following important equivalence results (originally derived by Färe and Zelenyuk (2012))), summarized in the next two corollaries.

Corollary 6. *We have*

$$\bar{e}_c(y^1, \dots, y^n, w) = \sum_{k=1}^n e_c(y^k, w) \times S^k, \quad (5.22)$$

where

$$S^k \equiv C^k(y^k, w) / \sum_{k=1}^n C^k(y^k, w). \quad (5.23)$$

This mathematical result is quite intuitive: In the dual framework, the scale elasticity of a group equals the *weighted* sum of the *individual* scale elasticity scores of all firms in this group, where the weights are the cost shares. As above, a strength of this result is that the weights are not *ad hoc* but derived from economic principles.

Similar aggregation result can also be derived for the *primal* scale elasticity measurement. In particular, from (5.11) and (5.18), we get the following equivalence result.

Corollary 7. *We have*

$$\bar{e}_i(y^1, \dots, y^n, x^*) = \sum_{k=1}^n e_i(y^k, x^{*k}) \times S^k. \quad (5.24)$$

This important result tells us how to obtain the *group* scale elasticity measure from the *individual* scale elasticity measures in the primal framework. Specifically, note that (5.24) says that one can get the primal aggregate scale elasticity measure from the *weighted* arithmetic average of the individual scale elasticity scores of all firms in this group, where the weights are the individual cost shares, derived from economic theoretic reasoning.

In case the researcher has no price information to calculate the weights, she/he may use shadow prices, estimated from the primal information or, alternatively, impose additional

assumption and help to derive the price independent weights, similarly as discussed above and following Färe and Zelenyuk (2003, 2007). Specifically, the additional assumption here would be

$$w_r \sum_{k=1}^n x_r^{*k} / \left(\sum_{r=1}^N w_r \sum_{k=1}^n x_r^{*k} \right) = b_r, \quad r = 1, \dots, N, \quad (5.25)$$

where $b_r \in (0, 1)$ is a known or estimated constant. In words, (5.25) states that the share of the group expenditures on the r^{th} input in the group total cost is given by b_r . Further, if we let $\varpi_r^k = x_r^k / \sum_{k=1}^n x_r^k$ be the share of the k^{th} firm in the group in terms of the r^{th} -input, then from (5.25) we get the price-independent weights given by

$$S^k = \sum_{r=1}^N \varpi_r^k b_r, \quad k = 1, \dots, n. \quad (5.26)$$

In words, (5.26) states that a firm's weight is the weighted average over all input-shares of this firm in the group, where the weights are the shares of the industry expenditures on the r^{th} input in the industry total cost.

It is also worth noting that analogous developments can also be done for other 'derivatives' of the cost function as well as of the revenue and profit functions. Moreover, such aggregation results can be generalized further to the case of aggregation *within sub-groups* (e.g., private vs. public, etc.) and then aggregation *between* these sub-groups into a larger group.

Finally, similar analysis can also be done for the case of aggregation of scale efficiency scores, as was done in Zelenyuk (2002, 2015).

6 Aggregation with Possibility of Reallocation

In the discussion above we restricted attention to cases where reallocation of inputs between DMUs are not allowed for the output orientation and reallocation of outputs between DMUs are not allowed for the input orientation. What if one of these or both restrictions are relaxed? This context was first considered in Nesterenko and Zelenyuk (2007) in the context of aggregating Farrell-type efficiency scores, while Mayer and Zelenyuk (2014a) extended it to the context of aggregating MPIS. Both papers focused on the output oriented context, while the input oriented context was outlined in Mayer and Zelenyuk (2014b) and refined further in Mayer and Zelenyuk (2017), which also extended it to the context of aggregating Hicks-Moorsteen Productivity Indexes (HMPIS). In this section we briefly summarize some key results from from these papers.

6.1 Aggregate Technology and Measures with Reallocation

To measure the gains from allowing for the reallocation of resources among DMUs in a group, we need to allow for a more general structure of aggregate technology, which we will refer to as the *group potential technology*, and define it as the Minkowski sum of technology sets of all individual DMUs for a given period τ .²¹

$$\Psi_\tau^* \equiv \sum_{\oplus k=1}^n \Psi_\tau^k. \quad (6.1)$$

While aggregating technology sets rather than the output or the input requirement sets (as was done above), this type of aggregate technology allows for full reallocation of inputs and outputs among all the DMUs in the group.²² Other, and equivalent, characterizations of this technology can be given via the *group potential input requirement* set, defined as

$$\mathcal{L}_\tau^*(\bar{Y}) = \{x : (x, \bar{Y}) \in \Psi_\tau^*\}, \quad (6.2)$$

and via the *group potential output* set, defined as

$$P_\tau^*(\bar{X}) = \{y : (\bar{X}, y) \in \Psi_\tau^*\}. \quad (6.3)$$

Based on this aggregate technology, and following Nesterenko and Zelenyuk (2007), let the *group potential output-oriented technical efficiency* be defined as²³

$$OTE_\tau^* \equiv OTE_\tau^*(\bar{X}_\tau, \bar{Y}_j) \equiv \sup_{\theta} \{\theta : \theta \bar{Y}_j \in P_\tau^*(\bar{X}_\tau)\}, \quad (6.4)$$

while the dual characterization of $P_\tau^*(\bar{X})$, the *group potential revenue function* is defined as

$$R_\tau^*(\bar{X}_\tau, p_j) \equiv \sup_y \{p_j y : y \in P_\tau^*(\bar{X})\}. \quad (6.5)$$

²¹This technology aggregation structure was earlier used in Li and Ng (1995), Blackorby and Russell (1999) and goes back to Koopmans (1957).

²²More recently, another definition of aggregate technology, which involved the union of technology sets, was considered by Peyrache (2013, 2015), which later was shown to be equivalent to the Koopmans-type aggregate technology Ψ_τ^* , under standard regularity conditions of production theory (see Zelenyuk (2018)).

²³Here, note that we allow for different time subscripts for inputs and outputs for the framework to be compatible with the HMPI context.

The associated *group potential revenue efficiency* is then defined as

$$RE_\tau^* \equiv RE_\tau^*(\bar{X}_\tau, \bar{Y}_j, p_j) \equiv \frac{R_\tau^*(\bar{X}_\tau, p_j)}{p_j \bar{Y}_j}, \quad p_j \bar{Y}_j \neq 0. \quad (6.6)$$

Due to duality between the revenue function and the output distance function, we have $RE_\tau^* \geq OTE_\tau^*$, and so the *group potential output-oriented allocative efficiency* can be defined to turn it into equality, yielding the following decomposition

$$RE_\tau^*(\bar{X}_\tau, \bar{Y}_j, p_j) = OTE_\tau^*(\bar{X}_\tau, \bar{Y}_j) \times OAE_\tau^*(\bar{X}_\tau, \bar{Y}_j, p_j), \quad \forall \tau, j. \quad (6.7)$$

In words, (6.4) and (6.6) measure the group efficiency relative to the group potential output set (6.3) and the associated aggregate cost function, similar to the individual level.

By the same token, and following Mayer and Zelenyuk (2014b, 2017), let the *group potential input-oriented technical efficiency* be defined as

$$ITE_\tau^* \equiv ITE_\tau^*(\bar{Y}_\tau, \bar{X}_j) \equiv \inf_\lambda \{ \lambda : \lambda \bar{X}_j \in \mathcal{L}_\tau^*(\bar{Y}_\tau) \}, \quad (6.8)$$

while the dual characterization of $\mathcal{L}_\tau^*(\bar{Y}_\tau)$, the *group potential cost function*, can be given by

$$C_\tau^*(\bar{Y}_\tau, w_j) \equiv \inf_x \{ w_j x : x \in \mathcal{L}_\tau^*(\bar{Y}_\tau) \}, \quad (6.9)$$

and so, the related *group potential cost efficiency* is then given by

$$CE_\tau^* \equiv CE_\tau^*(\bar{Y}_\tau, \bar{X}_j, w_j) \equiv \frac{C_\tau^*(\bar{Y}_\tau, w_j)}{w_j \bar{X}_j}, \quad w_j \bar{X}_j \neq 0. \quad (6.10)$$

As before, due to duality between the cost function and the input distance function, we have $CE_\tau^* \leq ITE_\tau^*$, and so the *group potential input-oriented allocative efficiency* can be defined to close this inequality, giving rise to the following decomposition

$$CE_\tau^*(\bar{Y}_\tau, \bar{X}_j, w_j) = ITE_\tau^*(\bar{Y}_\tau, \bar{X}_j) \times IAE_\tau^*(\bar{Y}_\tau, \bar{X}_j, w_j), \quad \forall \tau, j. \quad (6.11)$$

In words, (6.8) and (6.10) measure group efficiency relative to the group potential input requirement set (6.2) and associated aggregate cost function, in a way similar to measurements done at the individual level.

6.2 Reallocation vs. No Reallocation

A natural question at this stage is the following: What is the relationship between the group technology when the full reallocation is allowed to those we considered earlier (which did not allow for the full reallocation)? The following simple, yet important lemma clarifies this question.

Lemma 1. *We have*

$$\overline{P}_\tau(X_\tau) \subseteq P_\tau^*(\overline{X}_\tau), \quad (6.12)$$

and

$$\overline{\mathcal{L}}_\tau(Y_\tau) \subseteq \mathcal{L}_\tau^*(\overline{Y}_\tau). \quad (6.13)$$

A proof of (6.12) is relatively simple and can be found in Nesterenko and Zelenyuk (2007) and the proof of (6.13) is analogous (Mayer and Zelenyuk (2017)). This lemma affirms what is expected on an intuitive level: The aggregate technology characterizations where full reallocation across firms is allowed must always embrace, as a special case, the aggregate technology where the full reallocation (of outputs in the input oriented case and of inputs in the output oriented case) is not permitted.

As a result of this lemma, note that for any $(Y_\tau, \overline{Y}_\tau, w_j)$ we must have

$$C_\tau^*(\overline{Y}_\tau, w_j) \leq \overline{C}_\tau(Y_\tau, w_j), \quad (6.14)$$

and so for any $(Y_\tau, \overline{Y}_\tau, w_j)$, we must also have

$$CE_\tau^* \leq \overline{CE}_\tau. \quad (6.15)$$

Similarly, for any $(X_\tau, \overline{X}_\tau, p_j)$ we have

$$R_\tau^*(\overline{X}_\tau, p_j) \geq \overline{R}_\tau(X_\tau, p_j), \quad (6.16)$$

and so for any $(X_\tau, \overline{X}_\tau, p_j)$ we must also have

$$RE_\tau^* \geq \overline{RE}_\tau. \quad (6.17)$$

To measure the difference in efficiency between these different levels of aggregation, Nesterenko and Zelenyuk (2007) introduced the concept of reallocative efficiency. Specifically, in the output oriented context, we now also have the *group revenue reallocative efficiency*, as the

multiplicative residual that closes the inequality (6.17), i.e.,

$$RRE_\tau^* \equiv RRE_\tau^*(X_\tau, \bar{X}_\tau, p_j) \equiv R_\tau^*(\bar{X}_\tau, p_j) / \bar{R}_\tau(X_\tau, p_j), \quad (6.18)$$

and we obtain a useful decomposition of the group revenue efficiency

$$RE_\tau^* = \overline{RE}_\tau \times RRE_\tau^*. \quad (6.19)$$

Meanwhile, in the input oriented context (following Mayer and Zelenyuk (2017)), we have the *group cost reallocative efficiency*, as the multiplicative residual which closes the inequality (6.15), i.e.,

$$CRE_\tau^* \equiv CRE_\tau^*(Y_\tau, \bar{Y}_\tau, w_j) \equiv C_\tau^*(\bar{Y}_\tau, w_j) / \bar{C}_\tau(Y_\tau, w_j), \quad (6.20)$$

and so we obtain a useful decomposition of the group cost efficiency:

$$CE_\tau^* = \overline{CE}_\tau \times CRE_\tau^*. \quad (6.21)$$

What is even more interesting is that both RRE_τ^* and CRE_τ^* can be further decomposed, as outlined in the following lemmas.

Lemma 2. *We have*

$$RRE_\tau^* = OTRE_\tau^* \times OARE_\tau^*, \quad \forall \tau, \quad (6.22)$$

where group output-oriented technical reallocative efficiency is:

$$OTRE_\tau^* \equiv OTE_\tau^* / \overline{OTE}_\tau, \quad (6.23)$$

and group output-oriented allocative reallocative efficiency is:

$$OARE_\tau^* \equiv OAE_\tau^* / \overline{OAE}_\tau. \quad (6.24)$$

This result is from Nesterenko and Zelenyuk (2007) and its input oriented analogue (from Mayer and Zelenyuk (2014b, 2017)) is outlined next.

Lemma 3. *We have*

$$CRE_\tau^* = ITRE_\tau^* \times IARE_\tau^*, \quad \forall \tau, \quad (6.25)$$

where group input-oriented technical reallocative efficiency is

$$ITRE_\tau^* \equiv ITE_\tau^* / \overline{ITE}_\tau, \quad (6.26)$$

and group input-oriented allocative reallocative efficiency is

$$IARE_\tau^* \equiv IAE_\tau^*/\overline{IAE}_\tau. \quad (6.27)$$

In words, these two lemmas say that the reallocative efficiency measures characterize the difference for the group between individual efficiency in each DMU and the collective efficiency, where outputs are allowed to be reallocated among DMUs in the input orientation or inputs are allowed to be reallocated among DMUs in the output orientation.

6.3 Aggregate vs. Individual Reallocative Measures of Efficiency

What about the individual counterparts of the reallocative efficiency measures that appeared in the previous sub-section? Using the path ‘from aggregate to individual’, Nesterenko and Zelenyuk (2007) introduced the reallocative measures for individual output-oriented DMUs, which for $k = 1, \dots, n$ are given by

$$RRE_\tau^k \equiv RE_\tau^*/RE_\tau^k, \quad (6.28)$$

$$OTRE_\tau^k \equiv OTE_\tau^*/OTE_\tau^k, \quad (6.29)$$

$$OARE_\tau^k \equiv OAE_\tau^*/OAE_\tau^k, \quad (6.30)$$

and then established the relationship between individual and group reallocative measures, which we summarize in the next lemma.

Lemma 4. *We have*

$$RRE_\tau^* = \left(\sum_{k=1}^n (RRE_\tau^k(x_\tau^k, y_j^k, p_j))^{-1} \times S_j^k \right)^{-1}, \quad (6.31)$$

$$OTRE_\tau^* = \left(\sum_{k=1}^n (OTRE_\tau^k(x_\tau^k, y_j^k))^{-1} \times S_j^k \right)^{-1}, \quad (6.32)$$

$$OARE_\tau^* = \left(\sum_{k=1}^n (OARE_\tau^k(x_\tau^k, y_j^k, p_j))^{-1} \times S_{ae,\tau,j}^k \right)^{-1}, \quad (6.33)$$

where

$$S_j^k \equiv \frac{p_j y_j^k}{p_j \sum_{k=1}^n y_j^k}, \quad S_{ae,\tau,j}^k \equiv \frac{p(y_j^k \times OTE_\tau^k(x_j^k, y_j^k))}{p \sum_{k=1}^n (y_j^k \times OTE_\tau^k(x^k, y^k))}, \quad k = 1, \dots, n. \quad (6.34)$$

Moreover, note that combining these results with the decompositions we derived above,

we also get the following useful decompositions of group potential revenue efficiency

$$RE_\tau^* = \overline{OTE}_\tau \times \overline{OAE}_\tau \times OTRE_\tau^* \times OARE_\tau^*. \quad (6.35)$$

Furthermore, Mayer and Zelenyuk (2014b, 2017), using the same logic as Nesterenko and Zelenyuk (2007), defined the corresponding reallocative measures for individual input-oriented DMUs (for $k = 1, \dots, n$)

$$CRE_\tau^k \equiv CE_\tau^g / CE_\tau^k, \quad (6.36)$$

$$ITRE_\tau^k \equiv ITE_\tau^g / ITE_\tau^k, \quad (6.37)$$

$$IARE_\tau^k \equiv IAE_\tau^g / IAE_\tau^k. \quad (6.38)$$

and then established the relationship between the individual and the group reallocative measures, which we summarize in the following lemma.

Lemma 5. *We have*

$$CRE_\tau^* = \left(\sum_{k=1}^K (CRE_\tau^k(y_\tau^k, x_j^k, w_j))^{-1} \times W_j^k \right)^{-1}, \quad (6.39)$$

$$ITRE_\tau^* = \left(\sum_{k=1}^K (ITRE_\tau^k(y_\tau^k, x_j^k))^{-1} \times W_j^k \right)^{-1}, \quad (6.40)$$

$$IARE_\tau^* = \left(\sum_{k=1}^K (IARE_\tau^k(y_\tau^k, x_j^k, w_j))^{-1} \times W_{ae,\tau,j}^k \right)^{-1}, \quad (6.41)$$

where

$$W_j^k \equiv \frac{p_j y_j^k}{p_j \sum_{k=1}^n y_j^k}, \quad W_{ae,\tau,j}^k \equiv \frac{p_j (y_j^k OTE_\tau^k(x_j^k, y_j^k))}{p_j \sum_{k=1}^n (y_j^k OTE_\tau^k(x_j^k, y_j^k))}, \quad k = 1, \dots, n. \quad (6.42)$$

If we combine these results with the decompositions derived above, then we get the following decomposition of group potential cost efficiency:

$$CE_\tau^* = \overline{ITE}_\tau \times \overline{IAE}_\tau \times ITRE_\tau^* \times IARE_\tau^*. \quad (6.43)$$

These key results can be further used for extending various aggregation results discussed above to allow full reallocation: For MPI it was done in Mayer and Zelenyuk (2014a), while for HMPI it was done in Mayer and Zelenyuk (2017). Related extensions for the aggregation of scale efficiency and scale elasticity as well as for the directional distance functions are yet

to be developed, which presents a fruitful field of research for the near future.

7 Remarks on Estimation of Aggregate Scores

It is worth emphasizing here that our discussion so far was mainly theoretical, and we had not restricted our attention to any particular estimator. Indeed, the aggregation theories we summarized here are fairly general and can serve as a background for any suitable estimator, whether it is based on Data Envelopment Analysis (DEA), Stochastic Frontier Analysis (SFA), Free Disposal Hull (FDH) approach or another appropriate paradigm. These well established approaches can be used to estimate the individual efficiency scores which then can be aggregated to obtain the corresponding estimates of most of the aggregate scores we presented above.

Note however that the group potential measures are not calculated from the individual efficiency scores, but require calculation directly from the group potential technology. Yet, after imposing two extra assumptions we can recover these measures from the individual scores as well. These two assumptions are actually very common for many methods in productivity and efficiency analysis, especially in DEA. In particular, we can assume that (i) the technology set Ψ_τ^k is the same for all DMUs within each period and that (ii) it is also convex, then, following Li and Ng (1995) and Nesterenko and Zelenyuk (2007) we get:

$$\Psi_\tau^* = n\Psi_\tau, \quad \forall k = 1, \dots, n, \quad \forall \tau, \quad (7.1)$$

which in turn, for any period τ , gives:

$$P_\tau^*(\bar{X}_\tau) = nP_\tau(\tilde{x}_\tau), \quad (7.2)$$

where $\tilde{x}_j \equiv n^{-1} \sum_{k=1}^n x_j^k$, and

$$\mathcal{L}_\tau^*(\bar{Y}_\tau) = n\mathcal{L}_\tau(\tilde{y}_\tau), \quad (7.3)$$

where $\tilde{y}_\tau \equiv n^{-1} \sum_{k=1}^n y_\tau^k$. Intuitively, $P_\tau(\tilde{x}_\tau)$ and $\mathcal{L}_\tau(\tilde{y}_\tau)$ are, respectively, the output set and the input requirement set of the ‘average DMU’ for the sample (i.e., a hypothetical DMU whose input-output allocation is the average of input-output allocations in the sample, in period τ).

Therefore, the output-oriented group potential efficiencies can be obtained as the efficiency measures of the average DMU in the group, i.e., we have:

$$OTE_\tau^*(\bar{Y}_\tau, \bar{X}_j) = OTE_\tau(\tilde{x}_\tau, \tilde{y}_j), \quad (7.4)$$

$$RE_{\tau}^*(\bar{X}_{\tau}, \bar{Y}_j, p_j) = RE_{\tau}(\tilde{x}_{\tau}, \tilde{y}_j, p_j), \quad (7.5)$$

$$OAE_{\tau}^*(\bar{X}_{\tau}, \bar{Y}_j, p_j) = OAE_{\tau}(\tilde{x}_{\tau}, \tilde{y}_j, p_j) = RE_{\tau}(\tilde{x}_{\tau}, \tilde{y}_j, p_j)/OTE_{\tau}(\tilde{x}_{\tau}, \tilde{y}_j), \quad (7.6)$$

where OTE , RE and OAE are as defined in (3.5), (3.7) and (3.9) respectively, with superscript k dropped, and presented in the inter-temporal context.

Similarly, the input-oriented group potential efficiencies are the same as the efficiency measures of the average DMU in the group, i.e., we have:

$$ITE_{\tau}^*(\bar{Y}_{\tau}, \bar{X}_j) = ITE_{\tau}(\tilde{y}_{\tau}, \tilde{x}_j), \quad (7.7)$$

$$CE_{\tau}^*(\bar{Y}_{\tau}, \bar{X}_j, w_j) = CE_{\tau}(\tilde{y}_{\tau}, \tilde{x}_j, w_j), \quad (7.8)$$

$$IAE_{\tau}^*(\bar{Y}_{\tau}, \bar{X}_j, w_j) = IAE_{\tau}(\tilde{y}_{\tau}, \tilde{x}_j, w_j) = CE_{\tau}(\tilde{y}_{\tau}, \tilde{x}_j, w_j)/ITE_{\tau}(\tilde{y}_{\tau}, \tilde{x}_j), \quad (7.9)$$

where ITE , CE and IAE are as defined in (5.3), (5.5) and (5.7) respectively, which we cast in the inter-temporal context, with superscript k dropped.

It is also worth reminding that (7.4) and (7.7) are the versions of aggregate efficiency measures suggested (without the theoretical developments as summarized here) and advocated by Førsund and Hjalmarsson (1979).

8 Concluding Remarks

In this chapter we briefly summarized some of the key results on aggregation in productivity and efficiency analysis. We mostly focused on the output orientation and pointed out that most of the results can be extended or generalized to derive analogous results for the input orientation as well as to various other contexts.

Here we conclude by briefly mentioning a few interesting extensions and applications of these and other related aggregation results. First, the aggregation results for the directional distance functions were developed by Färe et al. (2008).²⁴ A similar theoretical framework for aggregating growth rates in the Solow's growth accounting approach was derived by Zelenyuk (2013a). Meanwhile, the theory for aggregation of the scale efficiency was developed by Zelenyuk (2015).

Related aggregation analysis can be also found in Pachkova (2009) and Raa (2011) and some interesting extensions can be found in Mussard and Peypoch (2006), Cooper et al. (2007), Li and Cheng (2007), Kuosmanen et al. (2010), Färe and Karagiannis (2014), Karagiannis (2015), Karagiannis and Lovell (2015).

²⁴Also see Zelenyuk (2002) for this and other related results.

Various applications analyzing real data for various economic questions can be found in Henderson and Zelenyuk (2007), Pilyavsky and Staat (2008), Weill (2008), Ferrier et al. (2009), Hall et al. (2012), Mugeru and Ojede (2014), to mention just a few.

Finally, in terms of actual estimation, note that what we discussed is a point-measure and one may (and typically should) be interested in the corresponding confidence interval measures and related inference. The first theoretical foundation for this important aspect was laid out in Simar and Zelenyuk (2007), who proposed a practical bootstrap-based approach for constructing confidence intervals and performing related inference on the aggregate efficiency measures. More recently, Simar and Zelenyuk (2018) extended this framework, by deriving convergence rates and new central limit theorems (CLTs) for the aggregate efficiency scores estimated via DEA and FDH. With the help of the Monte Carlo study, they also confirmed that for statistical inference on aggregate efficiency, the standard CLTs work poorly even for very simple 1-input-1-output cases and do not work correctly at all for larger dimensions. Meanwhile, the new CLTs that they derived performed reasonably well, reaching the nominal levels when samples get large. While deriving their asymptotic results, Simar and Zelenyuk (2018) focused on aggregates of the Farrell-type efficiency scores and similar developments are yet to be made for the other aggregates, which shall constitute key research questions in the area.

References

- Blackorby, C., Russell, R. R., 1999. Aggregation of efficiency indices. *Journal of Productivity Analysis* 12 (1), 5–20.
- Cooper, W., Huang, Z., Li, S., Parker, B., Pastor, J., 2007. Efficiency aggregation with enhanced Russell measures in data envelopment analysis. *Socio-Economic Planning Sciences* 41 (1), 1–21.
- Domar, E. D., 1961. On the measurement of technological change. *The Economic Journal* 71 (284), 709–729.
- Färe, R., Grosskopf, S., Lindgren, B., Roos, P., 1994. Productivity developments in Swedish hospitals: A Malmquist output index approach. In: Charnes, A., Cooper, W., Lewin, A. Y., Seiford, L. M. (Eds.), *Data Envelopment Analysis: Theory, Methodology and Applications*. Boston, MA: Kluwer Academic Publishers, pp. 253–272.
- Färe, R., Grosskopf, S., Lovell, C. A. K., 1986. Scale economies and duality. *Zeitschrift für Nationalökonomie / Journal of Economics* 46 (2), 175–182.
- Färe, R., Grosskopf, S., Zelenyuk, V., 2004a. Aggregation bias and its bounds in measuring technical efficiency. *Applied Economics Letters* 11 (10), 657–660.
URL <https://doi.org/10.1080/1350485042000207243>
- Färe, R., Grosskopf, S., Zelenyuk, V., 2004b. Aggregation of cost efficiency: Indicators and indexes across firms. *Academia Economic Papers* 32 (3), 395–411.
- Färe, R., Grosskopf, S., Zelenyuk, V., 2008. Aggregation of Nerlovian profit indicator. *Applied Economics Letters* 15 (11), 845–847.
- Färe, R., Karagiannis, G., 2014. A postscript on aggregate Farrell efficiencies. *European Journal of Operational Research* 233 (3), 784–786.
- Färe, R., Karagiannis, G., 2017. The denominator rule for share-weighting aggregation. *European Journal of Operational Research* 260 (3), 1175–1180.
URL <https://ideas.repec.org/a/eee/ejores/v260y2017i3p1175-1180.html>
- Färe, R., Primont, D., 1995. *Multi-Output Production and Duality: Theory and Applications*. New York, NY: Kluwer Academic Publishers.
- Färe, R., Zelenyuk, V., 2002. Input aggregation and technical efficiency. *Applied Economics Letters* 9 (10), 635–636.

- Färe, R., Zelenyuk, V., 2003. On aggregate Farrell efficiencies. *European Journal of Operational Research* 146 (3), 615–620.
- Färe, R., Zelenyuk, V., 2005. On Farrell’s decomposition and aggregation. *International Journal of Business and Economics* 4 (2), 167–171.
- Färe, R., Zelenyuk, V., 2007. Extending Färe and Zelenyuk (2003). *European Journal of Operational Research* 179 (2), 594–595.
- Färe, R., Zelenyuk, V., 2012. Aggregation of scale elasticities across firms. *Applied Economics Letters* 19 (16), 1593–1597.
- Färe, R., Zelenyuk, V., 2019. On luenberger input, output and productivity indicators. *Economics Letters* 179, 72 – 74.
- Farrell, M. J., 1957. The measurement of productive efficiency. *Jornal of the Royal Statistical Society. Series A (General)* 120 (3), 253–290.
- Ferrier, G., Leleu, H., Valdmanis, V., 2009. Hospital capacity in large urban areas: Is there enough in times of need? *Journal of Productivity Analysis* 32 (2), 103–117.
- Førsund, F. R., Hjalmarsson, L., 1979. Generalised Farrell measures of efficiency: An application to milk processing in Swedish dairy plants. *The Economic Journal* 89 (354), 294–315.
- Hall, M. J., Kenjegalievaa, K. A., Simper, R., 2012. Environmental factors affecting Hong Kong banking: A post-Asian financial crisis efficiency analysis. *Global Finance Journal* 23 (3), 184–201.
- Henderson, D. J., Zelenyuk, V., 2007. Testing for (efficiency) catching-up. *Southern Economic Journal* 73 (4), 1003–1019.
- Karagiannis, G., 2015. On structural and average technical efficiency. *Journal of Productivity Analysis* 43 (3), 259–267.
- Karagiannis, G., Lovell, C. A. K., 2015. Productivity measurement in radial DEA models with a single constant input. *European Journal of Operational Research* 251 (1), 323–328.
- Koopmans, T., 1957. *Three Essays on The State of Economic Science*. New York, NY: McGraw-Hill.

- Krein, M., Smulian, V., 1940. On regular convex sets in the space conjugate to a Banach space. *Annals of Mathematics* 41 (2), 556–583.
- Kuosmanen, T., Cherchye, L., Sipiläinen, T., 2006. The law of one price in data envelopment analysis: Restricting weight flexibility across firms. *European Journal of Operational Research* 170 (3), 735–757.
- Kuosmanen, T., Kortelainen, M., Sipiläinen, T., Cherchye, L., 2010. Firm and industry level profit efficiency analysis using absolute and uniform shadow prices. *European Journal of Operational Research* 202 (2), 584–594.
- Li, S. K., Cheng, Y. S., 2007. Solving the puzzles of structural efficiency. *European Journal of Operational Research* 180 (2), 713–722.
- Li, S.-K., Ng, Y. C., 1995. Measuring the productive efficiency of a group of firms. *International Advances in Economic Research* 1 (4), 377–390.
- Mayer, A., Zelenyuk, V., 2014a. Aggregation of Malmquist productivity indexes allowing for reallocation of resources. *European Journal of Operational Research* 238 (3), 774–785.
- Mayer, A., Zelenyuk, V., 2014b. An aggregation paradigm for Hicks-Moorsteen productivity indexes, cEPA Working Paper No. WP01/2014.
- Mayer, A., Zelenyuk, V., 2017. Aggregation of individual efficiency measures and productivity indices. In: Greene, W. H., ten Raa, T. (Eds.), *Handbook of Economic Performance Analysis*. Palgrave Macmillan, London, forthcoming.
- Mugera, A., Ojede, A., 2014. Technical efficiency in African agriculture: Is it catching up or lagging behind? *Journal of International Development* 26 (6), 779–795.
- Mussard, S., Peypoch, N., 2006. On multi-decomposition of the aggregate Malmquist productivity index. *Economics Letters* 91 (3), 436–443.
- Nesterenko, V., Zelenyuk, V., 2007. Measuring potential gains from reallocation of resources. *Journal of Productivity Analysis* 28 (1–2), 107–116.
- Oks, E., Sharir, M., 2006. Minkowski sums of monotone and general simple polygons. *Discrete and Computational Geometry* 35 (2), 223–240.
- Pachkova, E. V., 2009. Restricted reallocation of resources. *European Journal of Operational Research* 196 (3), 1049–1057.

- Panzar, J. C., Willig, R. D., 1977. Free entry and the sustainability of natural monopoly. *The Bell Journal of Economics* 8 (1), 1–22.
- Peyrache, A., 2013. Industry structural inefficiency and potential gains from mergers and break-ups: A comprehensive approach. *European Journal of Operational Research* 230 (2), 422–430.
- Peyrache, A., 2015. Cost constrained industry inefficiency. *European Journal of Operational Research* 247 (3), 996–1002.
- Pilyavsky, A., Staat, M., 2008. Efficiency and productivity change in Ukrainian health care. *Journal of Productivity Analysis* 29 (2), 143–154.
- Raa, T. T., 2011. Benchmarking and industry performance. *Journal of Productivity Analysis* 36 (3), 285–292.
- Schneider, R., 1993. *Convex Bodies: The Brunn-Minkowski Theory*. New York, NY: Cambridge University Press.
- Shephard, R. W., 1953. *Cost and Production Functions*. Princeton, NJ: Princeton University Press.
- Shephard, R. W., 1970. *Theory of Cost and Production Functions*. Princeton studies in mathematical economics. Princeton, NJ: Princeton University Press.
- Sickles, R. C., Zelenyuk, V., 2019. *Measurement of Productivity and Efficiency: Theory and Practice*. Cambridge University Press.
- Simar, L., Zelenyuk, V., 2007. Statistical inference for aggregates of Farrell-type efficiencies. *Journal of Applied Econometrics* 22 (7), 1367–1394.
- Simar, L., Zelenyuk, V., 2018. Central limit theorems for aggregate efficiency. *Operations Research* 166 (1), 139–149.
- Starr, R. M., 2008. Shapley-Folkman theorem. In: Durlauf, S. N., Blume, L. E. (Eds.), *The New Palgrave Dictionary of Economics*. Basingstoke, UK: Palgrave Macmillan, pp. 317–318.
- Tauer, L. W., 2001. Input aggregation and computed technical efficiency. *Applied Economics Letters* 8, 295–297.

- Weill, L., 2008. On the inefficiency of European socialist economies. *Journal of Productivity Analysis* 29 (2), 79–89.
- Wilson, P. W., 2018. Dimension reduction in nonparametric models of production. *European Journal of Operational Research* 267 (1), 349 – 367.
URL <http://www.sciencedirect.com/science/article/pii/S0377221717310317>
- Ylvinger, S., 2000. Industry performance and structural efficiency measures: Solutions to problems in firm models. *European Journal of Operational Research* 121 (1), 164–174.
- Zelenyuk, V., 2002. Essays in efficiency and productivity analysis of economic systems. Ph.D. thesis, Oregon State University, Corvallis, OR, USA.
- Zelenyuk, V., 2006. Aggregation of Malmquist productivity indexes. *European Journal of Operational Research* 174 (2), 1076–1086.
- Zelenyuk, V., 2013a. A scale elasticity measure for directional distance function and its dual: Theory and DEA estimation. *European Journal of Operational Research* 228 (3), 592–600.
- Zelenyuk, V., June 2013b. A note on equivalences in measuring returns to scale. *International Journal of Business and Economics* 12 (1), 85–89.
- Zelenyuk, V., 2015. Aggregation of scale efficiency. *European Journal of Operational Research* 240 (1), 269–277.
- Zelenyuk, V., 2018. Some mathematical and historical clarifications on aggregation in efficiency and productivity analysis and connection to economic theory, CEPA Working Paper No. WP03/2018.