Accepted Manuscript

Aggregation of scale efficiency

Valentin Zelenyuk

PII: S0377-2217(14)00541-4
DOI: http://dx.doi.org/10.1016/j.ejor.2014.06.038
Reference: EOR 12382

To appear in: *European Journal of Operational Research*

Received Date: 3 October 2013
Accepted Date: 25 June 2014

Please cite this article as: Zelenyuk, V., Aggregation of scale efficiency, *European Journal of Operational Research* (2014), doi: http://dx.doi.org/10.1016/j.ejor.2014.06.038

Aggregation of Scale Efficiency

Valentin Zelenyuk*

Last update: 14 March 2014

Abstract

In this article we generalize the aggregation theory in efficiency and productivity analysis by deriving solutions to the problem of aggregation of individual scale efficiency measures, primal and dual, into aggregate primal and dual scale efficiency measures of a group (e.g., industry). The new aggregation result is coherent with aggregation framework and solutions that were earlier derived for other related efficiency measures and can be used in practice for estimation of scale efficiency of an industry or other groups of firms within it.

Keywords: Scale Efficiency, Aggregation, Industry Efficiency, Duality, DEA.

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 thanks the editor and the anonymous referees for the very fruitful comments and suggestions—they helped improving the paper substantially.

1 Introduction

Analysis of economies of scale has been one of the fundamental subjects in economics, operations research and production management, both in theory and especially in practice. Indeed, the issue of attaining optimal scale of operations frequently appears among priority questions on management agendas of various types of companies, whether private or public. In this article we focus on the scale efficiency, which combines both the notion of optimal scale and the notion of (relative) production efficiency. Our particular interest is in how to appropriately measure the scale efficiency of *a group*—an industry consisting of firms, a firm consisting of plants, a union consisting of countries, etc.

For several decades, many studies in various economics and business literature challenged the issue of proper measurement of scale economies for various contexts and for various estimators. These studies include the seminal works of Hanoch (1975), Panzar and Willig (1977), Førsund and Hjalmarsson (1979), Banker (1984), Banker et al (1984), Färe and Grosskopf (1985), Färe, Grosskopf and Lovell (1986), Banker and Thrall (1992), Førsund (1996), Golany and Yu (1997), as well as more recent works of Førsund and Hjalmarsson (2004), Krivonozhko et al. (2004), Hadjicostas and Soteriou (2006, 2010), Podinovski et al. (2009), Zelenyuk (2013a,b,c), Peyrache (2013), to mention just a few.

In previous studies, researchers primarily focused on the measurement of scale economies for an *individual* or *disaggregate* decision making unit (DMU). In the present work we will focus on the issue of how to appropriately aggregate such individual scale efficiency measures, or their estimates/scores obtained from these measures for individual DMUs, into aggregate measures (or aggregate scores) of scale efficiency of a group. Indeed, after obtaining many individual scale efficiency scores, researchers may truly need a proper way to summarize these many scores into one or few numbers to present to their audience concisely. Clearly, one could simply use a sample average of individual estimates—but which one: the arithmetic or the geometric? More importantly, a problem with a sample average, whether arithmetic or geometric, is that it ignores a relative weight of each DMU in the aggregation. On the other hand, while a weighted average can account for a relative weight of each DMU, a critical question arises along the use of a weighted average: Which set of weights should be used? Indeed, conclusions and related policy implications may heavily depend on the weights chosen for the aggregation. Therefore, our primary focus is on deriving the weights for the aggregation of scale efficiency measures, which has not been done in the literature so far.

Studies on the aggregation problem in efficiency analysis go back to at least the seminal work of Farrell (1957), who coined the term structural efficiency of an industry. This notion was then criticized and elaborated by Førsund and Hjalmarsson (1979), who introduced the notion of efficiency of an average unit, and by Li and Ng (1995) who synthesized these latter works, considering them in the context of aggregation weights based on shadow prices. (See also a related discussion in Ylvinger (2000).) On the pure theoretical front, in their seminal work, Blackorby and Russell (1999) unveiled several impossibility theorems for a general aggregation problem of efficiency measures. One important implication of their work was that any positive result on aggregation in efficiency measurement must involve additional assumptions. This route was taken by Färe and Zelenyuk (2003) who, upon accepting certain assumptions on aggregate technology, optimization behavior and prices, and applying a revenue analogue of the fundamental theorem from Koopmans (1957), derived a solution to the aggregation problem for the output oriented Farrell-type technical efficiency measures. Similar approach was later used for deriving various aggregation results, such as aggregation of input oriented technical efficiency measures (Färe et al., 2004), aggregation of productivity indexes (Zelenyuk, 2006), aggregation within and between the sub-groups (Simar and Zelenyuk, 2007), aggregation of economic growth rates (Zelenyuk, 2011), but these works do not answer how the commonly used *scale* efficiency measures should be aggregated.

In the present work, we generalize the existing approach of aggregation of efficiency measures to the context of aggregation of scale efficiency measures, such that it is coherent with and encompass previous aggregation results for the related efficiency measures. This is a new theoretical result that can be relatively easily applied in practice for obtaining aggregate scale efficiency measures from suitable estimates of the individual scale efficiency scores.

The paper is structured as following: Section 2 briefly outlines the theory of measuring scale efficiency on individual level. Section 3 briefly outlines useful relationships between various efficiency measures that will be used in derivations of solutions to the aggregation problem. Section 4 proposes a solution to the aggregation problem from the perspective of mathematical (functional equations) approach. Section 5 outlines economic approach to solve the aggregation problem. Section 6 presents some special cases and Section 7 concludes.

2 Characterization of Individual Efficiencies

To keep our context general yet as simple as possible, let us consider a group of n decision making units (plants or firms or countries, etc.), hereafter DMUs, indexed by $k = 1, 2, \ldots, n$. Our main focus in the paper will be on aggregate efficiencies—measures that would represent various types of efficiency of a group of DMUs (e.g., a firm consisting of plants, an industry consisting of firms, etc.). We also want such aggregate measures of a group to be related or constructed from the individual efficiency measures obtained for each DMU in such a group. Following somewhat standard notation for defining individual efficiency measures, let $x^k = (x_1^k, \ldots, x_N^k)' \in \mathbb{R}_+^N$ be a vector of N inputs that a DMU k uses to produce a vector of M outputs, denoted by $y^k = (y_1^k, \ldots, y_M^k)' \in \mathbb{R}_+^M$. Furthermore, we assume that technology of a DMU $k \in \{1, 2, \ldots, n\}$ can be characterized by the technology set denoted by T^k and defined, in general terms, as

$$T^k = \left\{(x, y) \in \mathbb{R}_+^N \quad \mathbb{R}_+^M : \text{ DMU } k \text{ can produce } y \text{ from } x\right\}. \tag{2.1}$$

Equivalently, technology can be characterized via the output correspondence $P^k : \mathbb{R}_+^N \to 2^{\mathbb{R}_+^M}$ that assigns to each input vector $x \in \mathbb{R}_+^N$ a subset of all output vectors $y \in \mathbb{R}_+^M$ that can be produced with this particular x, and for a DMU $k \in \{1, 2, \ldots, n\}$ it is formally defined as

$$P^k(x) = \left\{y \in \mathbb{R}_+^M \ : \ (x, y) \in T^k\right\}, \quad x \in \mathbb{R}_+^N. \tag{2.2}$$

We assume that technology set satisfies standard regularity conditions of production theory (see Färe and Primont (1995)). To involve duality results, we also assume that all output sets, $P^k(x)$, are convex for all $x \in \mathbb{R}_+^N$, $k \in \{1, 2, \ldots, n\}$. This is a common assumption in economics, which is coherent with the principle of decreasing marginal rate of technical transformation regarding the substitution between outputs that is commonly assumed in microeconomics theory. Note, however, that for our general developments we do not assume convexity of T^k, although making such assumption is also possible (as is sometimes done in practical estimations), yet it would impose additional restrictions and so we stay free from it here.[1]

Using these characterizations of technologies, we will focus on the output orientation in efficiency measurement, particularly focusing on the following Farrell-type or radial output oriented measure of technical (in)efficiency, formally defined for a DMU $k \in \{1, 2, \ldots, n\}$ as

$$TE^k(x, y) = \sup\left\{\theta \in \mathbb{R}_{++} \ : \ \theta y \in P^k(x)\right\}, \quad (x, y) \in T^k. \tag{2.3}$$

Incidentally, note that for any $(x, y) \in T^k$, we have $TE^k(x, y) \geq 1$ and we will focus only on the practical case when $TE^k(x, y) < \infty$. Thus, $1/TE^k(x, y) \in (0, 1]$, i.e., the reciprocal of (2.3) gives a score between 0 and 1, with 1 standing for full or 100% technical efficiency level, output oriented.[2]

[1] Here, for the sake of space, we will focus on output orientation only, while the input orientation case can be derived in similar manner (in which case, instead of convexity of the output sets, one would need to assume convexity of the input requirement sets, defined as $L^k(y^k) = \left\{x \in \mathbb{R}_+^N \ : \ (x, y) \in T^k\right\}, \quad y \in \mathbb{R}_+^M$).

[2] Note there appears to be some confusion in the literature about this measure. Some authors use the output oriented Farrell technical efficiency measure defined as a reciprocal of (2.3), which in some cases might be preferred as it immediately has a convenient scale between 0 and 1. While the derivations below can be

Originally, Farrell (1957) focused on (an input analogue of) this measure for the context of a particular case of activity analysis model. Later, in their seminal paper, Charnes et al (1978) resurrected Farrell's approach and unveiled interesting dual characterization of his efficiency measure, with shadow price and shadow revenue-cost ratio interpretations, sparking a new area in operations research and econometrics and coining it as data envelopment analysis (DEA). While the DEA would be a natural estimator of (2.3), other estimators, such as regression-based methods, including stochastic frontier analysis (SFA), or a synthesis of them, such as stochastic-DEA method (e.g., see Simar and Zelenyuk, 2011) can work for our context as well. Therefore, we will consider our theoretical developments for a general case, regardless of what estimator is to be used for it. Also note that, besides not requiring convexity of T, unlike in the commonly assumed DEA context, we also allow for each DMU $k \in \{1, 2, \ldots, n\}$ to have its own unique technology (with possibility to be inefficient w.r.t. it), for the sake of generality. Clearly, assuming a common technology (as is often done in practice) would be a special case, and we briefly return to it in Section 6.

We will also use a dual characterization of the output correspondence and of efficiency measures—the revenue function, which for a DMU $k \in \{1, 2, \ldots, n\}$ with an input allocation $x \in \mathbb{R}^N_+$ is defined as

$$R^k(x, p) = \max_y \left\{ py \; : \; y \in P^k(x) \right\}, \quad x \in \mathbb{R}^N_+, \; p \in \mathbb{R}^M_+, \tag{2.4}$$

where $p = (p_1, ..., p_M) \in \mathbb{R}^M_+$ is the row-vector of prices corresponding to the (column) output vector y and here we will focus only on the practical case when $RE^k(x, p) \neq 0$.

To achieve our aggregation result, throughout this work we will focus on the output price vector p that is the same across all DMUs, which one could think of as a benchmark or a reference price vector selected for the purpose of aggregation. For example, if we consider a microeconomics context, then p could be taken to be a vector of equilibrium prices at the output markets. Alternatively, p can also be chosen to be the vector of shadow prices for the outputs for the considered group of firms (e.g., see Li and Ng (1995) for a motivation), etc. The differences in prices that one may face in practice can then be viewed as a variation or noise around these benchmark (equilibrium or shadow or other selected common) prices.[3]

Besides serving as an alternative (dual) complete characterization of technology of a DMU, due to duality theory in economics, the revenue function (2.4) also serves as the dual

written with this definition as well, we used definition (2.3) because it is more convenient for our derivations that follow, and may help avoiding confusions because previous works on aggregation of efficiency we refer to also used (2.3). To avoid confusions, we also refer to (2.3) as 'Farrell-type measure' rather than 'Farrell measure'. We thank anonymous referee for inspiring this comment.

[3] See Kuosmanen, Cherchye and Sipiläinen (2006) and Zelenyuk (2006, 2011) for a more extensive discussion about this assumption.

to the Farrell's output oriented technical efficiency measure (2.3), given regularity conditions, convexity of $P^k(x)$ and strictly positive output prices, and so it is often used to define the dual efficiency measures. Specifically, for a given DMU $k \in \{1, 2, \ldots, n\}$ with a combination (x, y, p), the revenue (in)efficiency measure that we will use is defined as the ratio of the maximal revenue implied for (x, p) to the actual revenue that the DMU k incurred for this same combination (x, p), i.e.,

$$RE^k(x, y, p) = \frac{R^k(x, p)}{py}, \text{ for } py \neq 0 \tag{2.5}$$

Incidentally, note that for any $(x, y) \in T^k$ and $p \in \mathbb{R}^M_{++}$ we have $1/RE^k(x, y, p) \in (0, 1]$, i.e., this measure of efficiency also gives a score between 0 and 1, with 1 standing for full or 100% revenue efficiency level.

Now, to define the scale efficiency measures, we will use an auxiliary, CRS-hypothetical technology defined, for each DMU $k \in \{1, 2, \ldots, n\}$, as

$$\check{T}^k = \left\{\delta(x, y) \; : \; (x, y) \in T^k, \; \forall \delta > 0\right\}, \tag{2.6}$$

i.e., $\check{T}^k$ is the set generated from T^k as the conical closure of T^k. Intuitively, $\check{T}^k$ can be understood as the smallest CRS technology set that includes the actual technology set T^k such that the upper boundary or technological frontier of $\check{T}^k$ is just tangent with that of T^k (at least at one point) and these tangent points are called the best possible scale allocations of (x, y).[4] For some technologies, such best scale allocations may be not unique as well as there might be uncountably infinite number of them. Also, it must be clear that, if an actual technology exhibits CRS (i.e., $T^k = \delta T^k, \forall \delta > 0$), then (and only then), by construction, we have $T^k = \check{T}^k$. This type of measurement is coherent with and frequently used in the DEA approach–in a sense it is just another, more general way to write what is usually done in DEA, so that the generalization allows for other methods to be used (e.g., SFA, stochastic-DEA, etc.).

Now, let us define the CRS-hypothetical output correspondence, $\check{P}^k : \mathbb{R}^N_+ \to 2^{\mathbb{R}^M_+}$ that assigns to each input vector $x \in \mathbb{R}^N_+$ the subset of all output vectors $y \in \mathbb{R}^M_+$ that can be produced with this particular x if technology was given by (2.6). Specifically, for a DMU $k \in \{1, 2, \ldots, n\}$, it is formally defined similarly to how we defined (2.2), as

$$\check{P}^k(x) = \left\{y \in \mathbb{R}^M_+ \; : \; (x, y) \in \check{T}^k\right\}, \quad x \in \mathbb{R}^N_+. \tag{2.7}$$

[4]Also see Frisch (1965) for discussion of the concept of technically optimal scale points.

So, by construction, $\check{P}^k(x)$ is an equivalent characterization of the CRS-hypothetical technology set $\check{T}^k$, in the sense that, for all $k \in \{1, 2, \dots, n\}$ we have

$$y \in \check{P}^k(x),\ x \in \mathbb{R}^N_+ \Longleftrightarrow (x, y) \in \check{T}^k. \tag{2.8}$$

Furthermore, the output oriented technical efficiency measure on the CRS-hypothetical technology for an allocation $(x, y) \in T^k$ for DMU $k \in \{1, 2, \dots, n\}$ is given by

$$\check{TE}^k(x, y) = \max \left\{ \lambda \in \mathbb{R}^1_{++} : \lambda y \in \check{P}^k(x) \right\}, \quad (x, y) \in T^k. \tag{2.9}$$

Note that for any $(x, y) \in T^k$ we have $\check{TE}^k(x, y) \geq 1$ and for technical purposes, to make the measurement of scale efficiency possible, we also focus on the case when $\check{TE}^k(x, y) < \infty, \forall k \in \{1, 2, \dots, n\}$. As a result we have $1/\check{TE}^k(x, y) \in (0, 1]$, i.e., this measure of efficiency gives a score between 0 and 1, with 1 standing for 100% technical efficiency level (output oriented) but w.r.t. the CRS-hypothetical technology rather than the original one.

Similarly, the revenue function with respect to the CRS-hypothetical technology for a DMU $k \in \{1, 2, \dots, n\}$ would be given by

$$\check{R}^k(x, p) = \max_y \left\{ py \ : \ y \in \check{P}^k(x) \right\}, \quad x \in \mathbb{R}^N_+,\ p \in \mathbb{R}^M_+, \tag{2.10}$$

while the associated revenue efficiency w.r.t. the CRS-hypothetical technology for a DMU $k \in \{1, 2, \dots, n\}$ facing allocation (x, y) and output prices p would be given by

$$\check{RE}^k(x, y, p) = \frac{\check{R}^k(x, p)}{py}, \quad \text{for } py \neq 0. \tag{2.11}$$

As it is for other efficiency measures defined above, for any $(x, y) \in T^k$ and $p \in \mathbb{R}^M_{++}$ we have $1/\check{RE}^k(x, y, p) \in (0, 1]$, i.e., the reciprocal of (2.11) gives a score between 0 and 1, with 1 standing for full or 100% revenue efficiency level w.r.t. the CRS-hypothetical technology.

Given (2.3) and (2.9) as well as (2.5) and (2.11), we can obtain primal and dual measures of scale efficiency. In particular, the primal (or technical or technology-based) scale efficiency measure for a DMU $k \in \{1, 2, \dots, n\}$ for an allocation $(x, y) \in T^k$ is defined as

$$TSE^k(x, y) = \frac{\check{TE}^k(x, y)}{TE^k(x, y)}, \tag{2.12}$$

Because $T^k \subseteq \check{T}^k$, for any $(x, y) \in T^k$ we have $\check{TE}^k(x, y) \geq TE^k(x, y)$, and therefore $1/TSE^k(x, y) \in (0, 1]$, $\forall k \in \{1, 2, \dots, n\}$, i.e., the reciprocal of this measure of efficiency gives a score between 0 and 1, with 1 standing for 100% primal scale efficiency level.

Similarly, the dual (or revenue based) scale efficiency measure for a DMU $k \in \{1, 2, \ldots, n\}$ with an input level x and output prices p, with $RE^k(x, p) \neq 0$, would be

$$RSE^k(x,p) = \frac{\check{RE}^k(x,y,p)}{RE^k(x,y,p)} = \frac{\check{R}^k(x,p)}{R^k(x,p)}, \quad (x,y) \in T^k,\ p \in \mathbb{R}^M_{++}. \tag{2.13}$$

Because $T^k \subseteq \check{T}^k$ we have $\check{R}^k(x,p) \geq R^k(x,p)$ and therefore $1/RSE^k(x,p) \in (0,1]$, $\forall k \in \{1, 2, \ldots, n\}$, i.e., the reciprocal of this measure of efficiency also gives a score between 0 and 1, with 1 standing for 100% revenue-based scale efficiency level. Note that $RSE^k(x,p)$ is independent from the output levels, which happens due to optimization behavior over outputs involved in (2.4) and (2.10) and due to benchmarking with respect to the same actual revenue py.

Our goal now is to find a proper way to aggregate the scores yielded by these *individual* scale efficiency measures, which we will derive with a help of important relationships and decompositions of the efficiency measures that we outline in the next section.

3 Key Decompositions of Individual Efficiency

In our further derivations we will utilize the following well-known decomposition of individual revenue efficiency measure, that hold for any $k \in \{1, 2, \ldots, n\}$,

$$RE^k(x,y,p) = TE^k(x,y) \times AE^k(x,y,p), \quad \forall (x,y) \in T^k,\ \forall p \in R^M_{++} \tag{3.1}$$

for the actual technology and, analogously for the CRS-hypothetical technology,

$$\check{RE}^k(x,y,p) = \check{TE}^k(x,y) \times \check{AE}^k(x,y,p), \quad \forall (x,y) \in T^k,\ \forall p \in R^M_{++} \tag{3.2}$$

where $AE^k(x,y,p)$ and $\check{AE}^k(x,y,p)$ are so-called allocative efficiency measures (output oriented), measuring inefficiency due to non-optimal (w.r.t. revenue optimization) allocation of outputs given p (see Färe and Primont (1995) for more details).

Because $\forall (x,y) \in T^k$ we have $\check{RE}^k(x,y,p) \geq \check{TE}^k(x,y)$ we have $1/\check{AE}^k(x,y,p) \in (0,1]$, $\forall k \in \{1, 2, \ldots, n\}$, i.e., the reciprocal of this measure of efficiency also gives a score between 0 and 1, with 1 standing for full or 100% of output oriented allocative efficiency level, but now w.r.t. the CRS-hypothetical technology and prices p.

Furthermore, note that we can also decompose $\check{RE}^k(x,y,p)$ into revenue efficiency and revenue based scale efficiency, as

$$\check{RE}^k(x,y,p) = RE^k(x,y,p) \times RSE^k(x,y,p), \quad (x,y) \in T^k,\ p \in \mathbb{R}^M_{++}. \tag{3.3}$$

Moreover, for any DMU $k \in \{1, 2, \ldots, n\}$, we can also decompose the individual revenue-based scale efficiency measure, $RSE^k(x, y, p)$, into technical and allocative parts, as

$$RSE^k(x,y,p) = TSE^k(x,y) \times ASE^k(x,y,p), \quad \forall (x,y) \in T^k, \forall p \in R^M_{++}. \tag{3.4}$$

where the latter component of (3.4) is given by

$$ASE^k(x,y,p) = \frac{\check{AE}^k(x,y,p)}{AE^k(x,y,p)}, \quad \forall (x,y) \in T^k, \forall p \in R^M_{++} \tag{3.5}$$

and we will refer to (3.5) as output oriented allocative scale (in)efficiency measure. Note, however, that unlike for revenue and technical efficiency measures, we cannot guarantee that $\check{AE}^k(x,y,p) \geq AE^k(x,y,p)$ or vice verse, and so one cannot guarantee in general that $1/ASE^k(x,y,p)$ is within $(0, 1]$.

Combining the above results, we get the following decomposition of the CRS-hypothetical revenue efficiency measure on the individual or disaggregate level

$$\check{RE}^k(x,y,p) = TE^k(x,y) \times AE^k(x,y,p) \times TSE^k(x,y) \times ASE^k(x,y,p), \tag{3.6}$$

and it holds for any $(x, y) \in T^k$, $p \in R^M_{++}$ and all DMUs $k \in \{1, 2, \ldots, n\}$.

In words, (3.7) decomposes the CRS-hypothetical-based revenue efficiency measure of a DMU k into four components: (i) the output oriented technical efficiency measure, (ii) the output oriented allocative efficiency measure, (iii) the output oriented technical scale efficiency measure, and (iv) the output oriented allocative scale efficiency measure. Now, the question is how to coherently aggregate all these measures from individual level into a group level, so that, preferably, such decomposition is also maintained at the aggregate level.

4 Efficiency Aggregation: Mathematical Approach

One way to describe the problem of aggregation we face here is by formulating a goal to find a sequence of aggregation functions $f_1, f_2, f_3, \ldots$ that would relate the aggregate efficiency measures, which we denote here with $\overline{\check{RE}}$, $\overline{RE}$, $\overline{TE}$, $\overline{AE}$, $\overline{RSE}$, $\overline{TSE}$ and $\overline{ASE}$, to the sets of their individual analogues. That is, we want to find some appropriate functions $f_1, f_2, f_3, f_4, f_5, f_6$ and f_7, where

$$\overline{\check{RE}} = f_1\left(\check{RE}^1, \ldots, \check{RE}^n\right) \tag{4.1}$$

$$\overline{RE} = f_2\left(RE^1, \ldots, RE^n\right) \tag{4.2}$$

$$\overline{TE} = f_3\left(TE^1, ..., TE^n\right) \tag{4.3}$$

$$\overline{AE} = f_4\left(AE^1, ..., AE^n\right) \tag{4.4}$$

$$\overline{RSE} = f_5\left(RSE^1, ..., RSE^n\right) \tag{4.5}$$

$$\overline{TSE} = f_6\left(TSE^1, ..., TSE^n\right) \tag{4.6}$$

$$\overline{ASE} = f_7\left(ASE^1, ..., ASE^n\right) \tag{4.7}$$

such that some desirable conditions on these relationships hold.

While our primal focus in this paper is on aggregation equations regarding the primal and dual measures of scale efficiency, i.e., on (4.5) and (4.6), it is natural to desire that our aggregation solutions to (4.5) and (4.6) are also coherent with solutions to aggregation problems regarding the other efficiency measures. For example, it is natural to desire that the functions $f_1, f_2, f_3, f_4, f_5, f_6$ and f_7 are such that we are able to obtain decompositions on the aggregate level that are analogous to those we can obtain on the individual level in (3.1), (3.3) and (3.4). That is, we may wish to require that the following relationships among the *aggregate* efficiency measures hold

$$\overline{RE} = \overline{TE} \times \overline{AE}, \tag{4.8}$$

$$\overline{\check{R}E} = \overline{RE} \times \overline{RSE}, \tag{4.9}$$

$$\overline{RSE} = \overline{TSE} \times \overline{ASE}. \tag{4.10}$$

In turn, these conditions (4.8)-(4.11) would also imply that we must have the following decomposition

$$\overline{\check{R}E} = \overline{TE} \quad \overline{AE} \quad \overline{TSE} \quad \overline{ASE}, \tag{4.11}$$

which is an analogue to the decomposition on the individual level given in (3.6).

Admitting such framework implies that the aggregation problem we face here is an example of a system of functional equations. If, in addition, we require that the weights of the aggregation (denoted here with ω^k, $k = 1, \ldots, n$) remain the same in all equations of the aggregation problem, then the only solution to this problem requires that all the aggregation functions are weighted geometric means (see Aczél (1990, p.27), Eichhorn (1978, p. 94) for more details). That is, the solution to the aggregation problem would be

$$\overline{\check{R}E} = \prod_{k=1}^{n} (\check{R}E^k)^{\omega^k} \tag{4.12}$$

$$\overline{RE} = \prod_{k=1}^{n} (RE^k)^{\omega^k} \tag{4.13}$$

$$\overline{TE} = \prod_{k=1}^{n} (TE^k)^{\omega^k} \tag{4.14}$$

$$\overline{AE} = \prod_{k=1}^{n} (AE^k)^{\omega^k} \tag{4.15}$$

$$\overline{TSE} = \prod_{k=1}^{n} (TSE)^{\omega^k} \tag{4.16}$$

$$\overline{ASE} = \prod_{k=1}^{n} (ASE^k)^{\omega^k} \tag{4.17}$$

$$\overline{RSE} = \prod_{k=1}^{n} (RSE^k)^{\omega^k}. \tag{4.18}$$

A weakness of such an approach, however, is that while formally answering the question about the aggregation function, this approach does not answer the question about what exactly the aggregation weights should be. A natural choice for the weights here in the output oriented or revenue-focused approach could be, for example, the observed revenue shares, i.e., $\omega^k = py^k / \sum_{k=1}^{n} py^k$, $k = 1, ...n$. This would be in the spirit of Farrell (1957) definition of structural efficiency of an industry. However, one could also argue that other weights (e.g., cost shares) might also be adequate. Moreover, there is no particular (economic) reason for the weights to be the same in all aggregation functions rather than being different for aggregating different efficiency measures. For example, some of the aggregation equations might have weights being the maximal (i.e., optimal w.r.t. p) revenue shares, given by $R^k(x,p) / \sum_{k=1}^{n} R^k(x,p)$, while other aggregation equations could use technically efficient shares for the aggregation, i.e., $py_*^k / \sum_{k=1}^{n} py_*^k$, where $y_*^k = y^k TE^k$ (e.g., see a related discussion on weights in Ylvinger (2000) and Färe and Zelenyuk (2007)). In fact, in the next sections, taking economic approach will indeed yield different weights for different efficiency measures.

It is worth noting here again that the issue of weights can often be even more important than the issue of aggregating functions—mainly because the aggregating functions are usually some type of averages and so usually yield similar results, especially for small variations of the scores being aggregated, while different weights can easily lead to dramatically different results and even may imply radically different policy implications. Our goal therefore is to derive some meaningful, economically compelling weights, as we do in the next section.

5 Efficiency Aggregation: Economics Approach

5.1 Aggregate Technology and Aggregate Efficiency

The goal of this sub-section is to define and outline characterization of *aggregate* technology. This aggregate technology will then be used to derive an aggregation scheme for aggregating scale efficiency such that it is coherent with aggregation of other related efficiency measures. To achieve this goal, let us denote the input and output allocations among DMUs within a group of interest by $X = (x^1, \ldots, x^n)$, which is an $N \times n$ matrix, and by $Y = (y^1, \ldots, y^n)$, which is an $M \times n$ matrix.

A critical step here is to define a group technology—the aggregate technology of all DMUs within the group. Following Färe and Zelenyuk (2003), one natural way for our (output oriented) context is to assume the additive structure of aggregation for the output sets

$$\overline{P}(X) = \overline{P}(x^1, ..., x^k, ..., x^n) = \sum_{k=1}^{n} P^k(x^k), \quad x^k \in R_+^N, \ k = 1, ..., n, \tag{5.1}$$

i.e., the aggregate output set is the Minkowski sum of the individual output sets.

The group or aggregate revenue function can now be obtained in a fashion similar to the definitions on the disaggregated level we had in (2.4), but now with respect to the aggregate technology given in (5.1), i.e.,

$$\overline{R}(X, p) = \max_{y} \left\{ py \ : \ y \in \overline{P}(X) \right\}, \quad p \in R_{++}^M, \ x^k \in R_+^N, \ k = 1, ..., n. \tag{5.2}$$

Similarly as was done on the disaggregate level, in (2.5), and letting $\overline{Y} = \sum_{k=1}^{n} y^k$ to denote the group's total actual output vector, we can use the aggregate revenue function to measure the aggregate revenue efficiency of a group that, facing output prices p, produced $\overline{Y}$ from an allocation of inputs X, via the following formula

$$\overline{RE}(X, \overline{Y}, p) = \frac{\overline{R}(X, p)}{p\overline{Y}}, \quad \text{for } p\overline{Y} \neq 0. \tag{5.3}$$

Intuitively, (5.3) is a measure of aggregate revenue efficiency of the group, that takes the maximal revenue obtainable from producing optimal (w.r.t. p) level of output and selling it at prices p, using the aggregate technology (5.1) and the matrix of input allocations X and compares it to the actual revenue of the group, given by $p\overline{Y}$. As other measures we considered above, $1/\overline{RE}(X, \overline{Y}, p)$ gives a score between 0 and 1, with 1 standing for full or 100% *aggregate* or group revenue efficiency level, w.r.t. the aggregate technology (5.1).

We now introduce new concepts that will help us deriving the new aggregation results. We start with the aggregate CRS-hypothetical technology, defined in a similar fashion as we did in (5.1), i.e.,

$$\check{\overline{P}}(X) = \check{\overline{P}}(x^1, ..., x^k, ..., x^n) = \sum_{k=1}^{n} \check{P}^k(x^k), \quad x^k \in R_+^N,\ k = 1, ..., n, \tag{5.4}$$

where $\check{P}^k(x^k)$ were defined in (2.7).

Intuitively, $\check{\overline{P}}(X)$ is the CRS-hypothetical analogue of $\overline{P}(X)$, constructed in the same fashion as the latter, but where the Minkowski summation is not over the original output sets but over their CRS-hypothetical analogues. Thus, $\check{\overline{P}}(X)$ will represent the aggregate CRS-hypothetical technology.

Based on this set-characterization of the aggregate CRS-hypothetical technology, we can define the aggregate dual (revenue-based) characterization of technology and of efficiencies, analogous to those we did on the individual level, in (2.10), (2.11) and (2.13). Specifically, and analogous to (2.10), the aggregate CRS-hypothetical revenue function can be obtained from

$$\check{\overline{R}}(X, p) = \max_y \left\{ py \ : \ y \in \check{\overline{P}}(X) \right\}, \quad p \in R_{++}^M,\ x^k \in R_+^N,\ k = 1, ..., n. \tag{5.5}$$

Intuitively, with this CRS-hypothetical aggregate revenue function (5.5), we look at what is the maximal level of total revenue the group can obtain by selling (at prices p) output produced from the endowed input allocations X (without their reallocation across DMUs) if the technology were given by $\check{\overline{P}}(X)$. Similarly as in (5.2), because the adopted aggregation structure (5.4) allows for reallocation of outputs but not for reallocation of inputs, the same nature of aggregation is then inherited by the aggregate CRS-hypothetical revenue function as well as by all the efficiency measures based on it.

Similarly to (2.11), the respective aggregate revenue efficiency w.r.t. the aggregate CRS-hypothetical technology can be measured by

$$\check{\overline{RE}}(X, \overline{Y}, p) = \frac{\check{\overline{R}}(X, p)}{p\overline{Y}}, \quad \text{for } p\overline{Y} \neq 0. \tag{5.6}$$

Intuitively, (5.6) is a measure of CRS-hypothetical aggregate revenue efficiency of the group that takes the maximal revenue obtained from selling (at prices p) output produced from X (without their reallocation across DMUs) and compares it to the actual revenues of the group, $p\overline{Y}$, assuming technology $\check{\overline{P}}(X)$ were feasible for this group. This latter proviso makes the entire difference between (5.6) and (5.3). Note also that $1/\check{\overline{RE}}(X, \overline{Y}, p)$ gives a score between 0 and 1, with 1 standing for full or 100% aggregate revenue efficiency level w.r.t. the aggregate CRS-hypothetical technology.

In its turn, we define the aggregate revenue scale efficiency measure of a group, analogously to those we have on the individual level, in (2.13),

$$\overline{RSE}(X,p) = \frac{\overline{\check{RE}}(X,\overline{Y},p)}{\overline{RE}(X,\overline{Y},p)} = \frac{\overline{\check{R}}(X,p)}{\overline{R}(X,p)}. \quad (5.7)$$

Note that $\overline{RSE}(X,p)$ is independent not only from Y and any individual y^k but also from the total outputs of the group, $\overline{Y}$, and total observed or actual revenues $p\overline{Y}$, which is due to the optimization (over outputs) involved in the construction of the revenue function as well as due to the process of benchmarking w.r.t. the same level of observed revenues, $p\overline{Y}$. Also note that because $\overline{\check{R}}(X,p) \geq \overline{R}(X,p)$, we have $1/\overline{\check{RSE}}(X,\overline{Y},p) \in (0,1]$, i.e., this measure also gives a score between 0 and 1, with 1 standing for 100% aggregate revenue scale efficiency level w.r.t. the aggregate CRS-hypothetical technology (5.4).

5.2 Individual vs. Aggregate Efficiency Measures

In this sub-section we establish relationships between the disaggregate (primal and dual) scale efficiency measures and their aggregate analogues under the aggregate technology structure defined in (5.1) and (5.4). In other words, we are interested in deriving aggregation functions and aggregating weights that relate the aggregate or group scale efficiency measures with their disaggregate or individual scale efficiency measures that are commonly estimated in practice. With the derivations that follow, we will justify the use of weighted arithmetic average aggregating function, where the weights (and the aggregation function) are derived from the economic optimization behavior and specific assumptions on aggregation of technologies and on output prices. Following Färe and Zelenyuk (2003) (also see Koopmans (1957)), we have

$$\overline{R}(X,p) = \sum_{k=1}^{n} R^k(x^k,p). \quad (5.8)$$

and therefore,

$$\overline{RE} = \overline{RE}(\overline{X},Y,w) = \sum_{k=1}^{n} RE^k(x^k,y^k,p)\omega^k, \quad (5.9)$$

where ω^k is the observed or actual (revenue-based) share-weight of DMU k in its group, i.e., formally

$$\omega^k = \frac{py^k}{\sum_{k=1}^{n} py^k},\ k = 1,...,n, \quad (5.10)$$

so that

$$\overline{RE} = \overline{TE} \times \overline{AE}, \tag{5.11}$$

where

$$\overline{TE} = \sum_{k=1}^{n} TE^k(x^k, y^k)\omega^k, \tag{5.12}$$

and

$$\overline{AE} = \sum_{k=1}^{n} AE^k(x^k, y^k, p)\omega_a^k, \tag{5.13}$$

with

$$\omega_a^k = \frac{py^k TE^k(x^k, y^k)}{\sum_{k=1}^{n} py^k TE^k(x^k, y^k)}. \tag{5.14}$$

Incidentally, note that the weights for aggregation of the allocative efficiency measures are different from those used for aggregation of the revenue and technical efficiency measures and, in particular, they account or correct for the technical inefficiency of the revenue shares used for weighting.

By the same line of proof as for (5.8) (see Färe and Zelenyuk (2003, 2007)), analogous result follows for the aggregate revenue function based on the aggregate CRS-hypothetical technology, i.e., we have

$$\overline{\check{R}}(X, p) = \sum_{k=1}^{n} \check{R}^k(x^k, p). \tag{5.15}$$

and

$$\overline{\check{RE}} = \overline{\check{RE}}(X, Y, p) = \sum_{k=1}^{n} \check{RE}(x^k, y^k, p)\omega^k. \tag{5.16}$$

where the individual weights ω^k, $k \in \{1, \ldots, n\}$ are the same as weights that appear in (5.9), and defined explicitly in (5.10).

In words, (5.16) says that the aggregate revenue efficiency measure w.r.t. the aggregate CRS-hypothetical technology, can be obtained by aggregating the individual revenue efficiency measures w.r.t. the individual CRS-hypothetical technology defined in (2.11), where the aggregating function is also the arithmetic average with individual weights given by the actual revenue shares of the individual k in the actual total revenue of its group.

Furthermore, note that from (5.7), it follows immediately that

$$\overline{\check{RE}}(X, \overline{Y}, p) = \overline{RE}(X, \overline{Y}, p) \quad \overline{RSE}(X, \overline{Y}, p). \tag{5.17}$$

i.e., we can decompose the aggregate revenue efficiency w.r.t. the aggregate CRS-hypothetical technology into two components: (i) the aggregate revenue efficiency (w.r.t. the aggregate technology) and (ii) the aggregate revenue scale efficiency.

More importantly, we can obtain the aggregate revenue scale efficiency of a group, $\overline{RSE}(X,\overline{Y},p)$, from the set of its individual analogues, $\left\{RSE^k(x^k,y^k,p)\right\}_{k=1}^n$. Specifically, combining (5.9) with (5.17) and applying some algebra, we get the following aggregation scheme:

$$\overline{RSE}(X,\overline{Y},p) = \sum_{k=1}^{n} RSE^k(x^k,y^k,p)\omega_r^k, \tag{5.18}$$

where

$$\omega_r^k = \frac{py^k RE^k(x^k,y^k,p)}{\sum_{k=1}^{n} py^k RE(x^k,y^k,p)} = \frac{R^k(x^k,p)}{\sum_{k=1}^{n} R^k(x^k,p)}. \tag{5.19}$$

Note that the weights that came out in the aggregation scheme (5.18), and described by (5.19), are different from those obtained for aggregation of revenue functions, technical and allocative efficiencies—here, they are the *efficient* revenue shares w.r.t. the individual revenue functions.

In turn, the characterization of the aggregate revenue scale efficiency given in (5.18)-(5.19), helps decomposing $\overline{\check{R}E}(X,\overline{Y},p)$ even further, by employing the decomposition of the aggregate revenue scale efficiency measure into technical and allocative parts. Specifically, after some algebra, we arrive to the following system of aggregation equations

$$\overline{RSE}(X,\overline{Y},p) = \overline{TSE} \times \overline{ASE} \tag{5.20}$$

where

$$\overline{TSE} = \sum_{k=1}^{n} TSE^k(x^k,y^k)\omega_r^k \tag{5.21}$$

and

$$\overline{ASE} = \sum_{k=1}^{n} ASE^k(x^k,y^k,p)\omega_{rts}^k \tag{5.22}$$

with ω_r^k given in (5.19), and

$$\omega_{rts}^k \quad = \quad \frac{R^k(x^k,p)TSE^k(x^k,y^k)}{\sum_{k=1}^{n} R^k(x^k,p)TSE^k(x^k,y^k)}. \tag{5.23}$$

Incidentally, note again that the weights that came out in the aggregation of the primal scale efficiency measures, in (5.21), are the same as those derived for the aggregation of the dual (revenue-based) scale efficiency measures, described by (5.19). On the other hand, the weights that come out in the aggregation of allocative scale efficiency measures, in (5.22), and described in (5.23) are different from those we derived earlier for the aggregation of revenue functions or for the aggregation of technical and allocative efficiency measures, as they also account for the primal scale inefficiency.

Finally, combining the statements above, we also get a desirable decomposition of the CRS-hypothetical revenue efficiency on the aggregate level—analogous to the decomposition we have on the disaggregated level that appeared in (3.6), i.e., we have

$$\overline{\check{RE}}(X, \overline{Y}, p) = \overline{TE} \times \overline{AE} \times \overline{TSE} \times \overline{ASE}. \tag{5.24}$$

In words, with the aggregation scheme derived from the revenue version of the Koopmans' theorem (and given our assumptions on the aggregate technology, same output prices, standard regularity conditions, etc.), we are able to decompose the *aggregate* CRS-hypothetical revenue efficiency measure into four components of different types of aggregate efficiency measures, namely: (i) the aggregate technical efficiency measure, (ii) the aggregate allocative efficiency measure, (iii) the aggregate primal scale efficiency measure, and (iv) the aggregate allocative scale efficiency measure. Incidentally, note that the product of the last two gives the aggregate dual scale efficiency, while the product of the first two gives the aggregate revenue efficiency measure. We thus attained the main goal of this paper—derived a coherent and economically compelling aggregation scheme for scale efficiency measures that embraces previous aggregation results, with a natural decomposition (5.24) that is analogous to decomposition (3.6) that exists at individual level. In the next sub-section we will discuss an important issue of proper stratification that is pertinent particularly to the context of aggregating scale efficiency measures.

5.3 Stratification

Various types of stratification of a group into sub-groups may be motivated for various empirical contexts, e.g., private vs. public firms in an industry, foreign vs. local firms, sub-groups of firms under different regulation regimes, sub-groups of countries in different economic unions or organizations (e.g., OECD vs. non-OECD countries, EU vs. non-EU, developed vs. developing countries, etc.) Accounting for such differences in a group and adapting aggregation scheme to such stratification may be vital in practice. The issue of efficiency aggregation within a sub-group and across several sub-groups into a larger group was first analyzed in Simar and Zelenyuk (2007) for the context of aggregation of technical and revenue efficiency measures. Similar approach can be applied here as well.

Importantly, even when there is no exogenous categorical variables that divide an analyzed group into distinct sub-groups, for the case of scale efficiency measurement there is a natural justification for stratifying a group of considered DMUs into potentially three sub-groups: (i) DMUs that are scale efficient, (ii) DMUs that are scale inefficient due to not exhausting the scale economies (i.e., too small to be scale efficient) and (iii) DMUs that are scale inefficient due to experiencing dis-economies of scale (i.e., too large to be scale efficient). Stratifying into these three groups, especially into the last two, and then proper aggregation within these sub-groups and between them is particularly important here because the reciprocals of the individual scale efficiency measures, whether primal or dual, give a score between 0 and 1 regardless of whether the firm is in group (ii) or in group (iii). In other words, the standard individual scale efficiency measures we considered above are not indicative of the source or cause of the scale inefficiency and ignoring this fact in producing the aggregate scale efficiency scores, whether primal or dual, would lead to missing important information about the scale economies. To do a proper aggregation, one should first stratify the estimates into the three groups and then apply the aggregation formulas derived above to aggregate the scale efficiency scores of DMUs within each of these three sub-groups separately. The resulting aggregate efficiency scores of sub-groups can then be aggregated into a larger group that consists of all or some of the sub-groups, using proper between-weights—the weights that would ensure hierarchical consistency of aggregation, as discussed in more details in Simar and Zelenyuk (2007).

6 Some Special Cases

It is now worth considering some interesting special cases for our aggregation problem. First of all, note that if all DMUs exhibit (or are to be measured with respect to) the same technology then the formulas derived above can still be applied without any changes.

Secondly, note that for all the types of efficiency measures considered here, except the allocative scale efficiency component, the aggregate efficiency is equal to one (i.e., 100% efficiency of the certain type) if and only if each individual efficiency score is equal to one.

Thirdly, note that each DMU in the group exhibits CRS technology if and only if

$$T\check{E}^k(x,y) = TE^k(x,y), \quad \forall k \in \{1,\ldots,n\}, \tag{6.1}$$

or, if and only if

$$\check{R}^k(x,p) = R^k(x,p), \; \forall k \in \{1,\ldots,n\}, \tag{6.2}$$

and therefore, if and only if

$$\overline{RSE} = \overline{TSE} = 1. \tag{6.3}$$

This is an important indication property of a scale efficiency measure that one may want an aggregate scale efficiency measure to satisfy.

Fourthly, if none of DMUs exhibits output oriented allocative inefficiency w.r.t. the original output set and w.r.t. its CRS-hypothetical technology (i.e., when $AE^k = \check{AE}^k = 1, \forall k \in \{1, \ldots, n\}$), then we have $\overline{ASE} = 1$ and therefore

$$\overline{RSE} = \overline{TSE} \tag{6.4}$$

which in turn will imply

$$\overline{\check{RE}} = \overline{TE} \times \overline{TSE}. \tag{6.5}$$

Although having no allocative inefficiency is a sufficient condition for the dual and primal (individual and aggregate) scale efficiencies to coincide, it is certainly not a necessary condition. Indeed, a much weaker condition for the dual and primal (individual and aggregate) scale efficiencies to yield equivalent results for any DMU k is for the technology of this DMU k to be output scale homothetic (see Zelenyuk (2013b)), i.e. to satisfy the following property

$$P^k(x) = H^k(x)\check{P}^k(x), \quad x \in \mathbb{R}^N_+. \tag{6.6}$$

where $H^k(x)$ is a finite real-valued semi-continuous function $H^k : \mathbb{R}^N_+ \to \mathbb{R}^1_+$ such that $H^k(x) \in (0,1], \forall y \in \mathbb{R}^M_+$. Intuitively, the structure of the individual technology of the type given by (6.6) assumes that the output set can be decomposed (in the multiplicative way) into the CRS-hypothetical output set, constructed from the original $P^k(x)$, and an appropriate scaling factor $H^k(x)$ that, in general, may depend only on the scale and the mix of inputs described by x.

With some algebra, it can then be shown (see Zelenyuk (2013b)) that a DMU k has output scale homothetic technology (6.6) if and only if

$$TE^k(x,y) = H^k(x)\check{TE}^k(x,y), \quad \forall (x,y) \in \mathbb{R}^{N+M}_+ \tag{6.7}$$

or, equivalently,

$$R^k(x,p) = H^k(x)\check{R}^k(x,p), \quad \forall x \in \mathbb{R}^N_+,\ p \in \mathbb{R}^M_{++}. \tag{6.8}$$

Therefore, an output scale homothetic technology is equivalent to saying that

$$TSE^k(x,y) = RSE^k(x,p) = 1/H^k(x), \quad \forall (x,y) \in \mathbb{R}_+^{N+M},\ p \in \mathbb{R}_{++}^M. \tag{6.9}$$

i.e., whenever the original output set differs from its CRS-hypothetical analogue by a scaling factor (the size of which may vary over the scale or mix of inputs). In other words, whenever $P^k(x)$ is a radial adjustment of $\check{P}^k(x)$ by some factor $H^k(x)$, we have exact equality of the primal and the dual individual scale efficiency measures.

In turn, if every DMU $k \in \{1, \ldots, n\}$ exhibits output scale homothetic technology as in (6.6), then

$$\overline{RSE} = \overline{TSE} = \sum_{k=1}^{n} (H^k(x^k))^{-1} \omega_r^k. \tag{6.10}$$

i.e., the aggregate primal and dual scale efficiency measures are also equal.

Finally, in case information on output prices needed to compute the aggregation weights is unavailable, additional assumptions can be imposed to derive price-independent weights (see Färe and Zelenyuk (2007) for details).

7 Concluding Remarks

In this paper we developed a theory for aggregation of scale efficiency measures across DMUs (firms, industries, countries, group of countries, etc.). The derivation is based on assuming optimization behavior, additive aggregation structure on the output sets and the same (e.g., equilibrium) output prices across all DMUs. An advantage of the resulting aggregation scheme (and aggregation weights, in particular) is that it is not ad hoc but derived from certain assumptions coherent with economic theory, and coherent with previous aggregation results.

This paper is just the first layer of theoretical foundation for analyzing aggregate scale efficiency measures. Its goal was to provide mathematical structure with economic theory foundation for such aggregation. The next layer shall be a statistical foundation. Indeed, besides presenting an average of results, researchers and practitioners are often interested in some measures of spread of the sample, such as the standard deviation, coefficient of variation, interquartile range, etc., as well as in the possibility to use some statistical testing procedures for inferring on various hypotheses. A hypothesis of an interest, for example, might be whether the true aggregate scale efficiency for a particular sub-group is different from unity (i.e., 100% efficiency) or some other level of interest. Another hypothesis of interest might be whether the aggregate scale efficiency scores are equal across different sub-groups or not, or across time for the same sub-groups, etc. Estimating characteristics of the sampling distribution of a statistic for a weighted mean where weights are random variables as well and related testing is a challenging task. A potential solution here would be to adapt the bootstrapping approach proposed in a different context by Efron (1979). This is a subject in itself, in some way similar to the recent work of Simar and Zelenyuk (2007) for bootstrapping the aggregate efficiency scores obtained via the DEA method, and we leave it for further research.

It is also worth noting that the derived weights (and the theory in general) may not be unique and perhaps better weight (and theories in general) could also be derived—and we hope this particular work will stimulate this to happen. A natural extension to the present work that might generate different weights would be to allow for reallocation of inputs between DMUs in the technology aggregation structure and one of the ways to do this is to adapt the approach of Nesterenko and Zelenyuk (2007) to the case of aggregation of scale efficiencies.

Another interesting research question is an extension of the presented theory to the intertemporal context—to measure changes of aggregate scale efficiency over time, as a component of aggregate Malmquist productivity index or other indexes, by extending the work of Zelenyuk (2006).

Similar aggregation theories can also be developed for the case when efficiency measures are based on other functions, e.g., hyperbolic distance function, the directional distance functions, etc. and these would be other natural paths for related future research (see Färe et al (2008) and Bobykin and Zelenyuk (2010) for some related results).

Overall, we believe that the aggregation solution with economic theory justification derived for the aggregation of scale efficiency in this paper must serve as a very useful platform for future applied as well as theoretical work.

References

[1] Aczél, J. (1990), "Determining Merged Relative Scores" Journal of Mathematical Analysis and Applications 150:1, 20–40.

[2] Banker, R.D., A.Charnes, W.W. Cooper (1984), "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis," Management Science 30:9, 1078-92.

[3] Banker, R.D. (1984) "Estimating most productive scale size using data envelopment analysis," European Journal of Operational Research 17, 35-44.

[4] Banker, R.D., and Thrall, R.M. (1992) "Estimation of returns to scale using Data Envelopment Analysis," European Journal of Operational Research 62, 74-84.

[5] Banker, R.D., Bardhan, I., Cooper, W.W. (1996) "A note on returns to scale in DEA" European Journal of Operational Research 88, 583-585.

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

[7] Bobykin, I. and V. Zelenyuk (2010) "Aggregation of Hyperbolic Efficiency" presented at North American Productivity Workshop (NAPW), at Rice University, Houston (USA), June 2010.

[8] Charnes, A., W.W. Cooper and E. Rhodes (1978), "Measuring the Efficiency of Decision Making Units," European Journal of Operational Research 2:6, 429-444.

[9] Efron, B. (1979), "Bootstrap methods: another look at the jackknife," Annals of Statistics 7, 1-26.

[10] Eichhorn, W. (1978), Functional Equations in Economics, Addison-Wesley, Reading, MA.

[11] Farrell, M. (1957), "The Measurement of Productive Efficiency," Journal of the Royal Statistical Society, Series A, General, 1 20, Part 3, 253-28.

[12] Färe, R. and S. Grosskopf (1985) "A Nonparametric Cost Approach to Scale Efficiency," Scandinavian Journal of Economics 87:4, 594-604.

[13] Färe, R. and S. Grosskopf (2000) "Notes on Some Inequalities in Economics," Economic Theory, 227-223.

[14] Färe, R., S. Grosskopf and C.A.K. Lovell (1986), "Scale Economies and Duality" Zeitschrift für Nationalökonomie 46:2, 175-182.

[15] Färe, R., S. Grosskopf and V. Zelenyuk (2008) "Aggregation of Nerlovian Profit Indicator," Applied Economics Letters 15, 845–847.

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

[17] Färe, R. and V. Zelenyuk (2012) "Aggregation of scale elasticities across firms," Applied Economics Letters 19, 1593-1597.

[18] Färe, R. and V. Zelenyuk (2007) "Extending Färe and Zelenyuk (2003)" European Journal of Operational Research 179:2.

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

[20] Førsund, F.R. (1996) "On the calculation of the scale elasticity in DEA models," Journal of Productivity Analysis 7, 283-302.

[21] Førsund, F.R. and L. Hjalmarsson (2004) "Calculating scale elasticity in DEA models," Journal of the Operational Research Society 55, 1023–1038.

[22] Førsund, F. R., and L. Hjalmarsson (1979) "Generalised Farrell measures of efficiency: An application to milk processing in Swedish dairy plants," The Economic Journal 89, 294–315.

[23] Frisch (1965), Theory of Production. Dordrecht: D. Reidel Publishing Co.

[24] Golany, B. and Yu, G. (1997) "Estimating returns to scale in DEA," European Journal of Operational Research 103, 28-37.

[25] Hadjicostas, P., and Soteriou, A.C. (2006) "One-sided elasticities and technical efficiency in multi-output production: a theoretical framework." European Journal of Operational Research 168, 425-449.

[26] Hadjicostas, P., and Soteriou, A.C. (2010) "Different orders of one-sided elasticities in multi-ouput production," Journal of Productivity Analysis 33, 147-167. Erratum (2010), 34, 181.

[27] Hanoch, G. (1975) "The Elasticity of Scale and the Shape of Average Costs," American Economic Review 65, 492-497.

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

[29] Kuosmanen, T., L. Cherchye, and T. Sipiläinen (2006) "The Law of One Price in Data Envelopment Analysis: Restricting Weight Flexibility Across Firms," European Journal of Operational Research 170:3, 735-757.

[30] Li, S.K., and Y.C. Ng (1995), "Measuring the Productive Efficiency of a Group of Firms," International Advances in Economic Research 1:4, 377-90.

[31] Nesterenko, V. and V. Zelenyuk (2007) "Measuring Potential Gains from Reallocation of Resources," Journal of Productivity Analysis 28:1, 107-116.

[32] Panzar, J.C. and R.D. Willig (1977) "Economies of Scale in Multi-Output Production, Quarterly Journal of Economics 91, 481-493.

[33] 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.

[34] Podinovski, V.V., Førsund, F.R., and Krivonozhko, V.E. (2009) "A simple derivation of scale elasticities in data envelopment analysis," European Journal of Operational Research 197, 149-153.

[35] Simar and Zelenyuk (2007) "Statistical Inference for Aggregates of Farrell-type Efficiencies" Journal of Applied Econometrics 22:7, 1367-1394.

[36] Simar, L. and V. Zelenyuk (2011) Stochastic FDH/DEA estimators for Frontier Analysis. Journal of Productivity Analysis 36:1, pp. 1-20.

[37] Ylvinger S., (2000) "Industry performance and structural efficiency measures: Solutions to problems in firm models," European Journal of Operational Research 121:1, 164-174.

[38] Zelenyuk, V. (2006) "On Aggregation of Malmquist Productivity Indexes," 2006, European Journal of Operational Research 174:2, 1076-1086.

[39] Zelenyuk, V. (2011) "Aggregation of Economic Growth Rates and of its Sources," European Journal of Operational Research 212, 190–198.

[40] 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, pp. 592-600.

[41] Zelenyuk, V. (2013b) "Scale efficiency and homotheticity: equivalence of primal and dual measures" (2013) Journal of Productivity Analysis, forthcoming.

[42] Zelenyuk, V. (2013c) "A note on equivalences in measuring returns to scale" International Journal of Business and Economics 12:1, pp. 85-89.

Research Highlights
for the article titled:
Aggregation of Scale Efficiency

- We extend the aggregation theory in efficiency and productivity analysis
- We derive solutions for aggregation of individual scale efficiency measures
- We provide practical way of estimating scale efficiency of an industry (e.g., from DEA).
- The new aggregation result is coherent with previous aggregation frameworks