Centre for Efficiency and Productivity Analysis

Working Paper Series
No. WP09/2011

Scale Efficiency: Equivalence of Primal and Dual Measures

Valentin Zelenyuk

Date:
December 2011

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

ISSN No. 1932 - 4398

Scale Efficiency: Equivalence of Primal and Dual Measures

Valentin Zelenyuk*

Abstract

In this paper we address an issue of equivalence of primal and dual measures of scale efficiency in general production theory framework. We find that particular types of homotheticity, which we refer to as *scale homotheticity*, provide necessary and sufficient condition for such equivalence. Specifically, we show that the input scale homotheticity of technology is necessary and sufficient condition for equivalence of primal and dual scale efficiency measures in the input/cost oriented case. Similarly, the output scale homotheticity of technology is necessary and sufficient condition for equivalence of primal and dual scale efficiency measures in the output/revenue oriented case. We also discuss the case when technology is both input scale homothetic and output scale homothetic, as well as indicate about some relationships of scale homotheticity with the homotheticity notions that have already been used in economic theory.

Keywords: Scale Efficiency, Homotheticity, Duality theory.

2 December 2011

Acknowledgement: I would like to thank Erwin Diewert, Rolf Färe, John Farrell, Shawna Grosskopf, Knox Lovell, Antonio Peyrache and Prasada Rao for their valuable comments. These individuals, of course, are not responsible for the views expressed here.

*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;

1. Introduction

Data limitation in empirical studies usually stands out as a rule rather than an exception. In production studies, for example, data on *all* inputs used in production are often hard or even impossible to obtain for an independent researcher. With a help of duality theory in economics (e.g., due to Shephard, 1953, 1970), vital economic information about production technology can be retrieved, for example, from data on input prices that firms faced and output levels these firms produced and the resulted total costs these firms incurred. Such information sometimes is more readily available to researchers than the primal (i.e., input-output) data, enabling researchers to estimate firm's cost function, obtaining an alternative (and complete) dual characterization of production technology, given some regularity conditions. In other contexts, a researcher might only have information on output prices that firms faced, the levels of inputs they used and the resulted total revenues these firms obtained, and then the duality theory can be used to estimate firms' revenue functions that provide yet another complete characterization of production technology, under some regularity conditions.

While giving complete characterizations of technology, the primal and various dual approaches to measuring the same issue in production (e.g., such as scale economies) do not always yield the same results, yet may be equivalent under some conditions. For example, in general, given some regularity conditions, the dual scale elasticity measures (based on cost, revenue or profit functions) provide the same information about scale economies of a technology as the primal scale elasticity measures, if the latter are evaluated at certain optimal allocations (w.r.t. cost, revenue or profit functions).[1] While such conditions of optimality might be natural in the context when firms are assumed to be fully efficient, such conditions might be incoherent with the contexts where one wants to allow and to measure inefficiency of firms and its sources, as is the context of our study.

In the case of production efficiency theory, a methodology to retrieve useful technology information using the cost-output data instead of input-output data goes back to at least Färe and Grosskopf (1985). In their seminal paper, authors provided conditions for equivalence of the primal (based on input-output data) and dual (based on cost-output data) *scale efficiency* measures in the context of activity analysis models (also known as Data Envelopment Analysis or DEA estimator). This condition requires that the input allocative efficiency estimated under the constant returns to scale (CRS) assumption is equal to that estimated under the variable returns to scale (VRS) assumption. This same condition was also envisioned informally in Seitz (1970).

The main goal of this paper is to theoretically determine what type of technology, if any, is consistent with this or other conditions that ensure equivalence of the primal and dual scale efficiency measures. It is known in economic theory literature that various types of homotheticity conditions on technology help establishing useful relationships of different measures of efficiency and productivity (e.g., see Chambers and Färe (1994, 1998), Førsund et al. (2001), Färe and Li (2002), etc.). What we find in our context is that, also, particular types of homotheticity, which we call here as *scale homotheticity*, provide necessary and sufficient condition for equivalence relationships we seek.

The paper is structured as following: In section 2 we briefly outline theoretical framework of technology characterization and efficiency measurement in the input oriented case. In section 3 we outline one of the main results of the paper, applied to the input oriented case, showing that input scale homothetic technology provides necessary and sufficient condition for the equivalence of the primal and dual measures in the input/cost oriented approach. In section 4 we extend the result of previous section to the output oriented case, showing that output scale homothetic technology provides necessary and sufficient condition for the equivalence of the primal and dual measures in the output/revenue oriented approach. In this section we also discuss the case when technology is simultaneously input scale homothetic and output scale homothetic. Finally, section 5 provides some concluding remarks and, in particular, mentions several possible extensions of this work.

2. Theoretical Framework: Input Oriented Case

In order to obtain general results, we look at the issue from a general theoretical perspective using the framework sparked by Shephard (1953, 1970) and further elaborated in Färe, Grosskopf and Lovell (1994) and Färe and Primont (1995). In particular, we assume that a production technology of interest can be characterized via technology set T, defined in general terms as

$$T = \{(x, y) \in \mathbb{R}_+^N \times \mathbb{R}_+^M : \ y \ \text{ is producible from } x \}. \tag{2.1}$$

and, equivalently, via the input correspondence $L: \mathbb{R}_+^M \to 2^{\mathbb{R}_+^N}$ that assigns to each output vector $y \in \mathbb{R}_+^M$ the subset of all input vectors $x \in \mathbb{R}_+^N$ that can produce this particular y, i.e.,

$$L(y) = \{x \in \mathbb{R}_+^N : (x, y) \in T\}, \qquad y \in \mathbb{R}_+^M \ . \tag{2.2}$$

and, equivalently, via the output correspondence $P: \mathbb{R}_+^N \to 2^{\mathbb{R}_+^M}$ that assigns to each input vector $x \in \mathbb{R}_+^N$ the subset of all output vectors $y \in \mathbb{R}_+^M$ that are producible from this particular x, i.e.,

$$P(x) = \{y \in \mathbb{R}_+^M : (x,y) \in T\}, \quad x \in \mathbb{R}_+^N . \tag{2.3}$$

We assume that technology is regular, i.e., satisfies standard regularity conditions of production theory (e.g., see Färe and Primont, 1995). In particular, we assume:

(i) technology set T is closed;

(ii) for any finite $x \in \mathbb{R}_+^N$, the output set $P(x)$ is bounded;

(iii) there is "no free lunch", i.e., $(\mathbf{0}_N, y) \notin T,\ \forall y \geq \mathbf{0}_M$;

(iv) "producing nothing is possible", i.e., $(x, \mathbf{0}_M) \in T,\ \forall x \in \mathbb{R}_+^N$;

(v) free disposability of outputs and free disposability of inputs holds, i.e.,

$(x^o, y^o) \in T \Longrightarrow (x,y) \in T, \forall y \leqq y^o\ \ \forall x \geqq x^o,\ \forall k \in \{1, \dots, n\}$.

To conveniently characterize technology set T and to measure technical efficiency, we use an implicit function $F_i : \mathbb{R}_+^M \times \mathbb{R}_+^N \rightarrow \mathbb{R}_+^1 \cup \{+\infty\}$, defined as

$$F_i(y,x) = \inf\{\lambda \in \mathbb{R}_{++}^1 : \lambda x \in L(y)\}, \tag{2.4}$$

which is referred to as the *Farrell* (1957) *input oriented measure of technical efficiency*. Incidentally, note that for any $x \in L(y)$ we have $F_i(y,x) \in (0,1]$. That is, the measure of efficiency in (2.4) gives a score between 0 and 1, with 1 standing for full or 100% technical efficiency level (in the input oriented case). This function is a reciprocal of the Shephard's (1953) input distance function and, given standard regularity conditions about T, completely characterizes T and L, in the sense that,[2]

$$L(y) = \{x \in \mathbb{R}_+^N : F_i(y,x) \leq 1\}, \qquad y \in \mathbb{R}_+^M . \tag{2.5}$$

To facilitate the measurement of scale efficiency in a general theoretical context, let $\check{T}$ be the CRS-hypothetical technology defined as

$$\check{T} \coloneqq \{\delta(x, y) : \quad (x, y) \in T, \ \forall \delta > 0\}. \tag{2.6}$$

Intuitively, $\check{T}$ is a set generated on T as the *smallest* conical closure of T. Intuitively, $\check{T}$ can be understood as the smallest CRS-hypothetical (or CRS-counterfactual) technology set that includes the actual technology set T. Thus, the upper boundary or technological frontier of $\check{T}$ would be just tangent with that of T at least at one point and such tangency point(s) would be called the best-possible-scale allocations of (x, y). Note that for some technologies, such best scale allocations may be not unique as well as there might be uncountably infinite number of them. Clearly, if the actual technology exhibits CRS, i.e., $T = \delta T, \forall \delta > 0$, then, by construction, we have $T = \check{T}$. For technical purposes, to make such measurement of scale efficiency possible, we assume that $\mathbb{R}_+^{\mathrm{N+M}} \backslash \check{T} \neq \emptyset$. This type of measurement is coherent with the approach used in DEA, but also can be used with other approaches, such as stochastic frontier analysis, stochastic DEA, etc.

Given the auxiliary set $\check{T}$, we now can define its dissections in the input space, which we denote with $L(y|crs)$, defined similarly to (2.2), as

$$L(y|crs) \coloneqq \left\{x \in \mathbb{R}_+^N : \ (x, y) \in \check{T}\right\}, \qquad y \in \mathbb{R}_+^M. \tag{2.7}$$

So, by construction, $L(y|crs)$ is an equivalent characterization of the CRS-hypothetical technology set $\check{T}$, in the sense that we have

$$x \in L(y|crs), \ \ y \in \mathbb{R}_+^M \quad \Leftrightarrow \quad (x, y) \in \check{T}. \tag{2.8}$$

We can now obtain functional characterizations of CRS-hypothetical technology and of efficiencies analogous to how we did in (2.4). Specifically, the input oriented Farrell measure of technical efficiency with respect to the CRS-hypothetical technology for an allocation (x, y) would be given by

$$F_i(y, x|crs) = \{\lambda \in \mathbb{R}^1_{++} : \ \lambda x \in L(y|crs)\}, \quad y \in \mathbb{R}^M_+. \tag{2.9}$$

Note that for any $x \in L(y)$ we also have $F_i(y, x|crs) \in (0,1]$, i.e., this measure of efficiency also gives a score between 0 and 1, with 1 standing for full or 100% technical efficiency level, but now w.r.t. the CRS-hypothetical technology rather than the original technology.

We can now use the technical efficiency measures in (2.4) and (2.9) to obtain the *input oriented scale efficiency* measure, as[3]

$$S_i(y, x) = \frac{F_i(y,x|crs)}{F_i(y,x)}, \qquad x \in L(y), y \in \mathbb{R}^M_+ \tag{2.10}$$

Because $T \subseteq \check{T}$, then $\forall(x, y) \in T$ we also have $F_i(y, x|crs) \leq F_i(y, x)$ and so $S_i(y, x) \in (0,1]$, i.e., this measure of efficiency also gives a score between 0 and 1, with 1 standing for full or 100% *scale efficiency* level (for input oriented case).

The approach to efficiency measurement presented just above is based on the input-output data, (x, y), and is referred to as the primal approach. To outline a dual, cost-oriented approach, let w be a row-vector of strictly positive input prices corresponding to each element of x, and define the (minimal) *cost function* as

$$C(y,w) = \min_x \{wx : x \in L(y)\}, \; y \in \mathbb{R}^M_+ : L(y) \neq \emptyset, \; w \in \mathbb{R}^N_{++}. \tag{2.11}$$

With additional assumption that $L(y)$ is convex, the dual analogue of (2.5) is given by

$$L(y) = \left\{x : \frac{C(y,w)}{wx} \leq 1, \; \forall\, w \in \mathbb{R}^N_{++}\right\}, \quad y \in \mathbb{R}^M_+, \tag{2.12}$$

i.e., technology can be completely characterized by the cost function (2.11), provided that technology satisfies standard regularity conditions and $L(y)$ is convex (see Färe and Primont (1995) for details).

The cost function in (2.11) can now be used to obtain the *cost efficiency* measure, a dual analogue of (2.4), via

$$CE(x,y,w) = \frac{C(y,w)}{wx}, \qquad x \in L(y), \; y \in \mathbb{R}^M_+, \; w \in \mathbb{R}^N_{++}. \tag{2.13}$$

And, the *cost-scale efficiency* for any allocation $x \in L(y), y \in \mathbb{R}^M_+$, and prices $w \in \mathbb{R}^N_{++}$ can then be measured via

$$S_c(y,w) = \frac{CE(x,y,w|crs)}{CE(x,y,w)} = \frac{C(y,w|crs)}{C(y,w)}, \tag{2.14}$$

where $C(y,w|crs)$ is the cost function with respect to the CRS-hypothetical technology, i.e.,

$$C(y,w|crs) = \min_x \{wx : x \in L(y|crs)\}, \quad y \in \mathbb{R}^M_+, w \in \mathbb{R}^N_{++} \tag{2.15}$$

while the associated cost efficiency w.r.t. the CRS-hypothetical technology would be

$$CE(x,y,w|crs) = \frac{C(y,w|crs)}{wx}, \quad x \in L(y), \; y \in \mathbb{R}^M_+ \quad w \in \mathbb{R}^N_{++}. \tag{2.16}$$

As it is for the other efficiency measures defined above, for any $x \in L(y)$ we have $CE(x, y, w) \in (0,1]$ and $CE(x, y, w|crs) \in (0,1]$, and because $T \subseteq \check{T}$, then $\forall(x, y) \in T$ we also have $CE(x, y, w|crs) \leq CE(x, y, w)$ and, as a result, $S_c(y, w) \in (0,1]$. That is, all these cost-based efficiency measures also give scores between 0 and 1, with 1 standing for full or 100% cost efficiency level.

Given convexity of $L(y)$, the relationship between the dual and primal approaches is obtained through the duality theory between $F_i(y, x)$ and $C(y, w)$. In particular, the Mahler's inequality (see Färe and Grosskopf (2000)) tells us that

$$\frac{C(y,w)}{wx} \leq F_i(y, x), \qquad \forall x \in L(y), y \in \mathbb{R}_{+}^{M},\ w \in \mathbb{R}_{++}^{N}. \tag{2.17}$$

This inequality can be 'closed' by defining the *allocative efficiency* measure as a multiplicative 'residual' that turns (2.17) into equality, i.e.,

$$\frac{C(y,w)}{wx} = F_i(y, x) AE_i(y, x, w), \qquad x \in L(y), y \in \mathbb{R}_{+}^{M}, w \in \mathbb{R}_{++}^{N}. \tag{2.18}$$

Analogously, we also have

$$\frac{C(y,w|crs)}{wx} \leq F_i(y, x|crs), \qquad \forall x \in L(y), y \in \mathbb{R}_{+}^{M}, w \in \mathbb{R}_{++}^{N}. \tag{2.19}$$

and so one can define the *allocative efficiency* measure w.r.t. CRS-hypothetical technology, as a multiplicative 'residual' that turns (2.19) into equality, i.e.,

$$\frac{C(y,w|crs)}{wx} = F_i(y, x|crs) AE_i(y, x, w|crs), \quad x \in L(y), y \in \mathbb{R}_{+}^{M},\ w \in \mathbb{R}_{++}^{N}. \tag{2.20}$$

Clearly, because the two measures of scale efficiency, $S_i(y,x)$ and $S_c(y,w)$, are functions of different variables, in general they may yield different efficiency scores. There is, however, a special case when they are equal. Specifically, one can observe that from (2.10), (2.14), (2.18) and (2.20) it follows that

$$S_c(y,w) = S_i(y,x) \times ASE_i(y,x,w), \qquad \forall x \in L(y), y \in \mathbb{R}_{+}^{M}, w \in \mathbb{R}_{++}^{N}, \quad (2.21)$$

where $ASE_i(y,x,w)$ is the allocative scale efficiency measure, defined as

$$ASE_i(y,x,w) = \frac{AE_i(y,x,w|crs)}{AE_i(y,x,w)}, \qquad x \in L(y),\ y \in \mathbb{R}_{++}^{M}, w \in \mathbb{R}_{++}^{N}. \quad (2.22)$$

Therefore, for any allocation $x \in L(y), y \in \mathbb{R}_{+}^{M}, w \in \mathbb{R}_{++}^{N}$, we can say

$$S_i(y,x) \ = \ S_c(y,w) \ \ \textit{if and only if} \ \ AE_i(y,x,w|crs) \ = \ AE_i(y,x,w). \quad (2.23)$$

The result in (2.23) was earlier reached by Färe and Grosskopf (1985) in the context of DEA or activity analysis modeling setup. An early reference for an intuitive explanation of this case is also found in Seitz (1970).

A natural theoretical question therefore arises: 'What type of technology can ensure an equivalence of the dual and primal measures of scale efficiency?' In particular, what are the conditions or assumptions on technology one may need to impose to ensure that, for all possible allocations, we have $AE_i(y,x,w|crs) = AE_i(y,x,w)$. The formal answer to this question is given in the next section.

3. Input Scale Homotheticity and Input Oriented Scale Efficiency Measurement

Consider a technology that satisfies the condition that will be referred to here as the *input scale homotheticity* (ISH), formally defined as

$$L(y) \;=\; G(y)L(y|crs), \qquad y \in \mathbb{R}_+^M\,, \tag{3.1}$$

where $G(y)$ is a continuous function $G:\ \mathbb{R}_+^M \rightarrow\ \mathbb{R}_+^1$ such that $G(y) \geq 1,\ \forall y \in \mathbb{R}_+^M.$ Intuitively, the structure of technology of the type given by (3.1) postulates that the input requirement set can be (multiplicatively) decomposed into the CRS-hypothetical input requirement set, constructed from the original $L(y)$, and an appropriate scalar augmentation factor $G(y)$ that, in general, may depend only on the scale and the mix of outputs.

Note that ISH is related to the concept of *input homotheticity* (see Färe and Mitchell (1993), and Färe and Primont (1995)).[4] In particular, recall that the *input homotheticity* is defined as

$$L(y) \;=\; H(y)L(1_M), \qquad y \in \mathbb{R}_+^M \tag{3.2}$$

for some function $H:\ \mathbb{R}_+^M \rightarrow\ \mathbb{R}_+^1$ consistent with regularity conditions on $L(y)$. Now, note that for a single-output and multi-input case, ISH implies that

$$L(y) \;=\; G(y)\; y\; L(1|crs), \quad y \in \mathbb{R}_+^1. \tag{3.3}$$

Rearranging and applying definition of ISH to the r.h.s. of (3.3) and then letting $H(y) = G(y)\; y \;/\; G(1)$, yields

$$L(y) \;=\; H(y)L(1), \qquad y \in \mathbb{R}_+^1\,, \tag{3.4}$$

which implies that for any $y \in \mathbb{R}^1_+$, if technology is ISH then it is also input homothetic (while the reverse is not true).

Next, consider what implications such ISH has towards the relationship between the primal and dual measures of scale efficiency. For this, first, note that ISH implies that

$$
\begin{aligned}
F_i(y,x) &= \inf_\lambda\{\lambda : \ x\lambda \in L(y)\} && \text{(by definition)} \\
&= \inf_\lambda\{\lambda : \ x\lambda \in G(y)L(y|crs)\} && \text{(by ISH)} \\
&= G(y)\inf_\lambda\left\{\frac{\lambda}{G(y)} : \ x\left(\frac{\lambda}{G(y)}\right) \in L(y|crs)\right\} \\
&= G(y)\inf_{\hat{\lambda}}\left\{\hat{\lambda} : \ x\hat{\lambda} \in L(y|crs)\right\} && \text{(where } \hat{\lambda} = \lambda/G(y)\text{)} \\
&= \ G(y)F_i(y,x|crs). && (3.5)
\end{aligned}
$$

On the other hand, assume that $F_i(y,x) \ = \ G(y)F_i(y,x|crs)$, then

$$
\begin{aligned}
L(y) &= \ \{x : \ F_i(y,x) \le \ 1\} && \text{(by definition)} \\
&= \ \{x: G(y)F_i(y,x|crs) \ \le \ 1\} \\
&= \ \left\{x : \ F_i\left(y, \frac{x}{G(y)}\middle| crs\right) \le 1\right\} && \text{(due to homogeneity of } F_i\text{)} \\
&= \ G(y)\,\{\hat{x} : \ \ F_i(y,\hat{x}|crs) \ \le \ 1\} && \text{(where } \hat{x} = \ x/G(y)\text{)} \\
&= \ G(y)L(y|crs). && (3.6)
\end{aligned}
$$

Combining the last two conclusions, (3.5) and (3.6), implies that ISH is equivalent to saying that

$$F_i(y,x) \ = \ G(y)F_i(y,x|crs), \qquad \forall x \in L(y), \ y \in \mathbb{R}^M_+ \,. \qquad (3.7)$$

At the same time, note that ISH also implies that

$$C(y,w) = \min_x\{wx : \ x \ \in \ L(y)\} \qquad \text{(by definition)}$$

$$= \min_x\{wx : x \in G(y)L(y|crs)\} \qquad \text{(by ISH)}$$

$$= \min_x\left\{wx : \frac{x}{G(y)} \in L(y|crs)\right\}$$

$$= G(y)\min_{\hat{x}}\{w\hat{x} : \hat{x} \in L(y|crs)\} \qquad \text{(where } \hat{x} = x/G(y)\text{)}$$

$$= G(y)C(y,w|crs). \tag{3.8}$$

On the other hand, if we assume $C(y,w) = G(y)C(y,w|crs)$, as well as assume convexity of $L(y)$ for any $y \in \mathbb{R}_+^M$, then we have

$$L(y) = \{x : wx \geq C(y,w),\ \forall\, w \in \mathbb{R}_{++}^N\} \qquad \text{(due to (2.12))}$$

$$= \{x : wx \geq G(y)C(y,w|crs), \forall\, w \in \mathbb{R}_{++}^N\}$$

$$= G(y)\left\{\frac{x}{G(y)} : w\left(\frac{x}{G(y)}\right) \geq C(y,w|crs), \forall\, w \in \mathbb{R}_{++}^N\right\}$$

$$= G(y)\{\hat{x} : w\hat{x} \geq C(y,w|crs), \forall\, w \in \mathbb{R}_{++}^N\} \qquad \text{(where } \hat{x} = x/G(y)\text{)}$$

$$= G(y)\, L(y|crs). \tag{3.9}$$

Therefore, combining the last two conclusions implies that ISH is equivalent to saying that

$$C(y,w) = G(y)C(y,w|crs),\ \forall y \in \mathbb{R}_+^M,\ L(y) \neq \varnothing,\ w \in \mathbb{R}_{++}^N, \tag{3.10}$$

In turn, combining (3.7) and (3.10), implies that ISH is equivalent to

$$AE_i(y,x,w|crs) = AE_i(y,x,w),\ \ \forall x \in L(y), y \in \mathbb{R}_+^M, w \in \mathbb{R}_{++}^N. \tag{3.11}$$

In words, the equivalence of input oriented allocative efficiency estimated under the CRS-hypothetical technology assumption to that estimated allowing for variable returns to scale can happen *if and only if* technology is input scale homothetic. That is, we reached an answer to the research question of this paper for the input oriented context:

In the input oriented case, the equivalence of the dual and primal scale efficiency measures for all technologically feasible allocations, i.e., $S_i(y,x) = S_c(y,w)$, for all $x \in L(y)$ and $w \in \mathbb{R}^N_{++}$, can be achieved *if and only if* technology is input scale homothetic.

Incidentally, notice that in such a case, both scale efficiency measures are equal to the reciprocal of $G(y)$—the function reflecting the scale of the production activity of such technology. That is, altogether, for all $x \in L(y)$ and $w \in \mathbb{R}^N_{++}$ we have

$$S_i(y,x) = S_c(y,w) = \frac{1}{G(y)} \quad \Leftrightarrow \quad \text{technology is ISH}\,. \tag{3.12}$$

The input scale homotheticity condition on technology may seem to be fairly restrictive to assume, yet it is the minimal (i.e., the necessary and sufficient) condition that guarantees equality of the primal and dual measures of scale efficiency in the input oriented framework. Moreover, note that ISH technology is less restrictive than CRS technology and is a generalization of it, since any CRS technology is necessarily an ISH technology but the reverse is not true. Also, we note that any technology with one input (and any finite number of outputs) that satisfy the regularity conditions mentioned above is necessarily ISH, provided that $\mathbb{R}^{\mathrm{N+M}}_{+} \backslash \breve{T} \neq \emptyset$.

4. Scale Efficiency Measurement in Output and Revenue Oriented Case

In practice, many researchers, especially in macroeconomic studies, prefer taking the output or revenue-focused orientation for efficiency measurement. Thus, the goal of this section is to briefly outline similar developments for the output oriented case, which in turn will facilitate further discussion of the case of simultaneous input and output scale homotheticity. In the output/revenue oriented context of efficiency (and productivity) measurement, researchers often employ the so-called *Farrell output oriented measure of technical efficiency*, which is an implicit function $F_o : \mathbb{R}_+^N \times \mathbb{R}_+^M \rightarrow \mathbb{R}_+^1 \cup \{+\infty\}$, defined as

$$F_o(x, y) = \sup\{\theta \in \mathbb{R}_{++}^1 : \theta y \in P(x)\}. \tag{4.1}$$

The function (4.1) is a reciprocal of the Shephard's (1970) output distance function and, given standard regularity conditions on technology set T, completely characterizes T and P, in the sense that

$$P(x) = \{y \in \mathbb{R}_+^M : F_o(x, y) \leq 1\}, \qquad x \in \mathbb{R}_+^N. \tag{4.2}$$

Also note that for any $y \in P(x)$ we have $1/F_o(x, y) \in (0,1]$, i.e., the reciprocal of (4.2) gives a score between 0 and 1, with 1 standing for full or 100% technical efficiency level (in the output oriented case).

As for the input oriented analog given in (2.10), the measure in (4.1) is also often used to construct the primal measure of *scale efficiency* for the output oriented context, as

$$S_o(x, y) = \frac{F_o(x,y|crs)}{F_o(x,y)}, \qquad y \in P(x) \tag{4.3}$$

where $F_o(x, y|crs)$ is the output oriented Farrell measure of technical efficiency with respect to the CRS-hypothetical technology for an allocation (x, y), i.e.,

$$F_o(x, y|crs) = \sup\{\theta \in \mathbb{R}_{++}^1 : \theta y \in P(x|crs)\}, \quad x \in \mathbb{R}_+^N. \tag{4.4}$$

where $P(x|crs)$ is an equivalent characterization of the CRS-hypothetical technology set $\check{T}$, i.e.,

$$P(x|crs) \coloneqq \{y \in \mathbb{R}_+^M : (x, y) \in \check{T}\}, \qquad x \in \mathbb{R}_+^N. \tag{4.5}$$

Also note that for any $y \in P(x)$ we have $1/F_o(x, y|crs) \in (0,1]$, i.e., the reciprocal of (4.4) gives a score between 0 and 1, with 1 standing for full or 100% technical efficiency level (in output oriented case) w.r.t. the CRS-hypothetical technology. Moreover, because $T \subseteq \check{T}$, then $\forall (x, y) \in T$ we also have $F_o(x, y|crs) \geq F_o(x, y)$ and, therefore, $1/S_o(y, x) \in (0,1]$, i.e., the reciprocal of (4.5) also gives a score between 0 and 1, with 1 standing for full or 100% *scale efficiency* level (in the output oriented context).

Again, this approach is based on the input-output data, (x, y), and is referred to as the primal approach in output oriented case, and to outline its dual, revenue-oriented, approach, let p be a row-vector of strictly positive output prices corresponding to each element of y, and define the (maximal) *revenue function* as

$$R(x, p) = \max_y \{py : y \in P(x)\}, \; x \in \mathbb{R}^N_+ \;, \; p \in \mathbb{R}^M_{++}. \tag{4.6}$$

If in addition to the standard regularity conditions we also assume convexity of $P(x)$, then the dual analogue of (4.2) can be given via

$$P(x) = \left\{y : \frac{R(x,p)}{py} \geq 1 \;, \; \forall\, p \in \mathbb{R}^M_{++}\right\}, \qquad x \in \mathbb{R}^N_+ \tag{4.7}$$

i.e., technology set T can be completely characterized by the revenue function (4.6), provided that this technology satisfies standard regularity conditions and $P(x)$ is convex (see Färe and Primont (1995) for details). Note that we do not restrict T to be convex, neither here, nor in the previous section, and in this section we do not need convexity of $L(y)$, although in practice T is oftent assumed to be convex, which in turn would imply convexity of $P(x)$ and of $L(y)$.

The *revenue efficiency* measure, a dual analogue of (4.1), is then given by

$$RE(x,y,p) = \frac{R(x,p)}{py}, \qquad y \in P(x),\ x \in \mathbb{R}_{+}^{N}, p \in \mathbb{R}_{++}^{M}. \tag{4.8}$$

and the *revenue-scale efficiency* can be measured via

$$S_r(x,p) = \frac{RE(x,y,p|crs)}{RE(x,y,p)} = \frac{R(x,p|crs)}{R(x,p)}, \quad y \in P(x), x \in \mathbb{R}_{+}^{N},\ p \in \mathbb{R}_{++}^{M}. \tag{4.9}$$

where $R(x,p|crs)$ is the revenue function with respect to the CRS-hypothetical technology, i.e.,

$$R(x,p|crs) = \max_y \{py : y \in P(x|crs)\},\ x \in \mathbb{R}_{+}^{N}, x \in \mathbb{R}_{+}^{N},\ p \in \mathbb{R}_{++}^{M}, \tag{4.10}$$

while the associated revenue efficiency w.r.t. the CRS-hypothetical technology would be

$$RE(x,y,p|crs) = \frac{R(x,p|crs)}{py}, \qquad y \in P(x),\ x \in \mathbb{R}_{+}^{N},\ p \in \mathbb{R}_{++}^{M}. \tag{4.11}$$

As for the other efficiency measures we considered above, note that for any $y \in P(x)$ we have $\left(RE(x,y,p|crs)\right)^{-1} \in (0,1]$, i.e., the reciprocal of (4.11) also gives a score between 0 and 1, with 1 standing for full or 100% revenue efficiency level w.r.t. the CRS-hypothetical technology (for output oriented case). Moreover, because $T \subseteq \check{T}$, then $\forall (x,y) \in T$ we have $RE(x,y,p|crs) \geq RE(x,y,p)$ and, as a result, $1/S_r(x,p) \in (0,1]$. That is, this measure of efficiency also gives a score between 0 and 1, with 1 standing for full or 100% *revenue-based scale efficiency* level (for output oriented case).

Given convexity of $P(x)$, the relationship between the output oriented technical efficiency and revenue efficiency is obtained through the duality theory between $F_o(x, y)$ and $R(x, p)$. In particular, the Mahler's inequality (Färe and Grosskopf (2000)) states that

$$\frac{R(x,p)}{py} \geq F_o(x, y), \qquad y \in P(x), x \in \mathbb{R}^N_+,\ p \in \mathbb{R}^M_{++}. \tag{4.12}$$

And, similarly as in (2.13), the inequality (4.12) is 'closed' by defining the output oriented *allocative efficiency* measure as a multiplicative 'residual' that converts (4.12) into equality, i.e.,

$$\frac{R(x,p)}{py} = F_o(x, y) AE_o(x, y, p), \qquad y \in P(x), x \in \mathbb{R}^N_+, p \in \mathbb{R}^M_{++}. \tag{4.13}$$

Furthermore, we also have

$$\frac{R(x,p|crs)}{py} \geq F_o(x, y|crs), \qquad y \in P(x), x \in \mathbb{R}^N_+,\ p \in \mathbb{R}^M_{++}. \tag{4.14}$$

and we define the output oriented *allocative efficiency* measure w.r.t. CRS-hypothetical technology as a multiplicative 'residual' that turns (4.14) into equality, i.e.,

$$\frac{R(x,p|crs)}{py} = F_o(x, y|crs) AE_o(x, y, p|crs), \quad y \in P(x),\ x \in \mathbb{R}^N_+,\ p \in \mathbb{R}^M_{++}. \tag{4.15}$$

Obviously, these two measures of scale efficiency, $S_o(x, y)$ and $S_r(x, p)$, in general may yield different efficiency scores, and similarly as for the input oriented case, we can always say that

$$S_r(x, p) = S_o(x, y) \times ASE_o(x, y, p), \quad y \in P(x),\ x \in \mathbb{R}^N_+,\ p \in \mathbb{R}^M_{++} \tag{4.16}$$

where we will refer to $ASE_o(x, y, p)$ as the allocative scale efficiency measure, defined as

$$ASE_o(x,y,p) = \frac{AE_o(x,y,p|C)}{AE_o(x,y,p)}, \qquad y \in P(x),\ x \in \mathbb{R}^N_+,\ p \in \mathbb{R}^M_{++}. \qquad (4.17)$$

And, therefore, for any $y \in P(x),\ x \in \mathbb{R}^N_+,\ p \in \mathbb{R}^M_{++}$, we have

$$S_o(x,y) = S_r(x,p) \quad \textit{if and only if} \quad AE_o(x,y,p|crs) = AE_o(x,y,p). \qquad (4.18)$$

Now, let us consider a technology that satisfy the condition that will be referred to here as the *output scale homotheticity* (OSH), defined as

$$P(x) = G_o(x)P(x|crs), \qquad x \in \mathbb{R}^N_+, \qquad (4.19)$$

where $G_o(x)$ is a continuous function $G_o : \mathbb{R}^N_+ \to \mathbb{R}^1_+$ such that $0 < G(y) \leq 1,\ \forall x \in \mathbb{R}^N_+$. Intuitively, (4.19) says that technology must be such that the output correspondence can be (multiplicatively) decomposed into the CRS-hypothetical output correspondence, constructed from the original $P(x)$ and an appropriate scalar augmentation factor $G_o(x)$ that, in general, may depend only on the scale and mix of inputs.

It is also worth noting that, similarly as ISH is related to input homotheticity, OSH is related to the concept of *output homotheticity* (Färe and Mitchell (1993), and Färe and Primont (1995)), defined as

$$P(x) = H_o(x)P(1_N), \qquad x \in \mathbb{R}^N_+, \qquad (4.20)$$

for some function $H_o : \mathbb{R}^N_+ \to \mathbb{R}^1_+$ consistent with regularity conditions on $P(x)$. To see this, note that for a single-input and multi-output case, OSH implies that

$$P(x) = G_o(x)\, x\, P(1|crs), \quad x \in \mathbb{R}^1_+. \qquad (4.21)$$

Rearranging and applying definition of OSH to r.h.s. of (4.21) and letting $H_o(x) = G_o(x)x/G_o(1)$, yields

$$P(x) = H_o(x)P(1), \qquad x \in \mathbb{R}^1_+, \tag{4.22}$$

implying that for any $x \in \mathbb{R}^1_+$, if technology is OSH then it is also output homothetic.

Now, to see the implications of such type of technology on the relationship between the primal and dual measures of scale efficiency in the output oriented case, note that OSH is equivalent to

$$F_o(x,y) = G_o(x)F_o(x,y|crs), \qquad x \in \mathbb{R}^N_+,\ p \in \mathbb{R}^M_{++}. \tag{4.23}$$

and, with assumption that $P(x)$ is convex, for any $x \in \mathbb{R}^N_+$, OSH is also equivalent to saying that

$$R(x,p) = G_o(x)R(x,p|crs), \qquad x \in \mathbb{R}^N_+,\ p \in \mathbb{R}^M_{++}, \qquad (4.24)^5$$

In turn, combining (4.23) and (4.24), implies that OSH is equivalent to

$$AE_o(x,y,p|crs) = AE_o(x,y,p), \quad \forall y \in P(x),\ p \in \mathbb{R}^M_{++}. \tag{4.25}$$

In words, the equivalence of output oriented allocative efficiency estimated under the CRS assumption to that estimated under the VRS assumption can happen *if and only if* the technology is output scale homothetic. In other words, we reach an answer to the research question of this paper for the revenue/output oriented case, which states that:

> The equivalence of the dual and primal scale efficiency measures for all technologically feasible allocations, in the output/ revenue oriented case, i.e.,
>
> $$S_o(x,y) = S_r(x,p), \text{ for all } y \in P(x) \text{ and } p \in \mathbb{R}^M_{++},$$

can be achieved *if and only if* technology is output scale homothetic. In particular, both measures of scale efficiency would yield $\frac{1}{G_o(x)}$, for all $y \in P(x)$ and $p \in \mathbb{R}^M_{++}$.

Similarly as the ISH, the output scale homotheticity condition on technology is also fairly restrictive, yet it is the minimal (i.e., the necessary and sufficient) condition that guarantees equality of the primal and dual measures of scale efficiency for any allocation, in the output oriented framework. Moreover, OSH technology is also less restrictive than CRS technology, generalizing it in the sense that any CRS technology is necessarily an OSH technology but the reverse is not always true. Furthermore, a frequent case in many applications is the case when technology has one output and some finite number of inputs, and such technologies are OSH, provided that $\mathbb{R}^{N+M}_{+} \backslash \check{T} \neq \emptyset$.

From the discussion above, a natural question arises: “What happens when a technology satisfies both to ISH and OSH?” For example, recall that simultaneous or joint input and output homotheticity is equivalent to the so-called inverse homotheticity (e.g., see Shephard (1970), Färe and Mitchell (1993) and Färe and Primont (1995)). Can we say something similar about simultaneous input and output scale homotheticity? To see some implications, first, note that $\check{T}$ exhibits CRS if and only if

$$F_o(x, y|crs) = \frac{1}{F_i(y,x|crs)}, \qquad \forall (x, y) \in \check{T}. \tag{4.26}$$

Therefore, combining (4.26) with (3.7) and (4.23), yields

$$F_o(x, y) = \frac{G(y) G_o(x)}{F_i(y,x)}, \qquad \forall (x, y) \in T. \tag{4.27}$$

That is, under the joint input and output homotheticity, we get a generalization of the well-known relationship stated in (4.26), where the output oriented technical efficiency can be found as the reciprocal of the input oriented technical efficiency, adjusted by the scaling factors $G(y)G_o(x)$. It is also worth noting that any regular technology with one input and one output is simultaneousely ISH and OSH. This particular case is certainly very restrictive for practical applications, yet it might be important to remember as many graphical illustrations (e.g., of productivity indexes, etc.) are often made for such 1-inpu-1-output examples.

Finally, it might be also worth noting that in general, for all allocations $(x, y) \in T$, the two primal measures of scale efficiency, $S_i(y, x)$ and $\left(S_o(x, y)\right)^{-1}$, are guaranteed to yield the same score if and only if technology set T is a cone, i.e., technology exhibits CRS. Of course, when the measurement of scale efficiency is made at points (x, y) that are simultaneously on the isoquant of $P(x)$ and on the isoquant of $L(y)$, i.e., where $F_o(x, y) = F_i(y, x) = 1$, then the two primal measures of scale efficiency, $S_i(y, x)$ and $\left(S_o(x, y)\right)^{-1}$, would yield identical scores. This result is analogous to the conclusion reached for the derivatives-based scale elasticity measures (e.g., see Färe, Grosskopf and Lovell (1986) and Zelenyuk (2011a, 2011b)), where under some regularity conditions, the input and output oriented measures also yield equivalent results, when they are measured at points where $F_o(x, y) = F_i(y, x) = 1$.

5. Concluding Remarks

In this paper we show that the equivalence of primal and dual scale efficiency measures hold *if and only if* technology is of a peculiar type—*scale homothetic*. Specifically, we had shown that input scale homotheticity is necessary and sufficient condition for equivalence of primal and dual scale efficiency measures in the input oriented case, while the output scale homotheticity is necessary and sufficient condition for equivalence of primal and dual scale efficiency measures in the output oriented case. We also discuss the case when technology is both input scale homothetic and output scale homothetic.

Our work also opens several directions for further research. One direction is to develop statistical tests for identifying ISH or OSH technology, e.g., by using similar developments as for other types of homotheticity (e.g., in Färe and Li (2002)) as well as recent developments on applications of non-parametric statistical methods in efficiency analysis.[6] Another direction is to make further theoretical investigations of how restrictive are the assumptions of ISH or OSH, or both, on technology, and what are their relationships to various types of Hicks neutrality and separability, e.g., as those investigated and catalogued by Chambers and Färe (1994) and Førsund et al. (2001).

It might be also useful to look at applications of ISH and OSH technology to other related areas. For example, a natural direction would be the area of productivity indexes, where the total productivity change is often decomposed into change in technology and change in efficiency and the latter component is often decomposed into change in 'pure-technical efficiency' and change in 'scale efficiency'. Thus, naturally, the (input or output) scale homotheticity condition can then help relating the primal and dual measures of scale efficiency change. Moreover, ISH or OSH, or both simultaneously, can be used to derive new relationships between the input oriented Malmquist productivity index and output oriented Malmquist productivity indexes, to simplify the Hicks-Moorsteen productivity index to depend only on the input oriented or output oriented distance functions, etc. Overall, we hope that this article will stimulate work on these and other interesting directions.

References

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

Chambers, R. and R. Färe (1994) "Hicks' Neutrality and Trade-Biased Growth: A Taxonomy," *Journal of Economic Theory*, pp. 554-567.

Chambers, R. and R. Färe (1998) "Translation Homotheticity," *Economic Theory*, pp. 629-641.

Färe, R. and F.R. Førsund, S. Grosskopf, K. Hayes, and A. Heshmati (2001) "A Note on Decomposing the Malmquist Productivity Index by Means of Subvector Homotheticity," *Economic Theory* pp. 239-245.

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

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

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

Färe, R. and S. Grosskopf and C.A.K. Lovell, (1994) *Production Frontier*, Cambridge, Cambridge University Press.

Färe, R. and T. Mitchell, (1993) "Multiple Outputs and Homotheticity," *Southern Economic Journal*, pp.287-296.

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

Färe, R. and S.K. Li (2001), "A Nonparametric Test of Translation Homotheticity," *Economics Letters* 72, pp. 341-345.

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

Førsund, F. and L. Hjalmarsson (1979), "Generalized Farrell measures of efficiency: an application to milk processing in Swedish dairy plants," *Economic Journal* 89, pp. 294-315.

Førsund, F. "On the Calculation of the Scale Elasticity in DEA Models" *Journal of Productivity Analysis* 7 (2-3), 283-302.

Shephard, R. W. (1953) *Cost and Production Functions*, Princeton: Princeton University Press.

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

Seitz, W.D. (1970), "The Measurement of Efficiency Relative to a Frontier Production Function," *American Journal of Agricultural Economics* 52, pp. 505-511.

Simar, L. and P.W. Wilson (2008), "Statistical Inference in Nonparametric Frontier Models: recent Developments and Perspectives," in *The Measurement of Productive Efficiency,* by Harold Fried, C. A. Knox Lovell and Shelton Schmidt, editors, Oxford University Press.

Simar, L. and P.W. Wilson (2010), Estimation and Inference in Cross-Sectional Stochastic Frontier Models, Econometric Reviews 29, 62-98.

Simar, L. and P.W. Wilson (2011), Inference by the m out of n bootstrap in nonparametric frontier models. *Journal of Productivity Analysis* 36, 33-53.

Simar, L. and V. Zelenyuk (2006) "On Testing Equality of Two Distribution Functions of Efficiency Score Estimated via DEA," *Econometric Reviews*, 25:4, 497-522.

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

Zelenyuk, V. (2002). "Essays in Efficiency and Productivity Analysis of Economic systems." PhD thesis. Corvallis: Oregon State University.

Zelenyuk, V. (2011a) "A Note on Equivalences in Measuring Returns to Scale in Multi-output-multi-input Technologies" Working Paper Series No. WP05/2011, Centre for Efficiency and Productivity Analysis, School of Economics, University of Queensland, Australia.

Zelenyuk, V. (2011b) "A Scale Elasticity Measure for Directional Distance Function and its Dual" Working Paper Series No. WP05/2011, Centre for Efficiency and Productivity Analysis, School of Economics, University of Queensland, Australia.

Notes

[1] See Färe, Grosskopf and Lovell (1986) and Färe and Primont (1995).

[2] For the definition of the Shephard's distance function, regularity conditions and the proof of this and other properties, see Färe and Primont (1995).

[3] The origins of this measure go back to at least Førsund and Hjalmarsson (1979). For this and other ways of measuring *scale* issues see, for example, Banker et al. (1984), Färe and Grosskopf (1985), Färe, Grosskopf and Lovell (1994), Førsund (1997), to mention just a few.

[4] I thank Rolf Färe for this insightful comment.

[5] We skip the derivations here for the sake of space as they are similar to those we did for the input oriented case.

[6] E.g., see Simar and Wilson (2008, 2010, 2011) and Simar and Zelenyuk (2006, 2007, 2011).