The Royal Veterinary and Agricultural University
Food and Resource Economic Institute

KVL A UNIVERSITY OF LIFE SCIENCES

Unit of Economics Working Papers 1999/7

Allocative Efficiency of Technically Inefficient Production Units.

Peter Bogetoft and
Rolf Färe

Allocative Efficiency of Technically Inefficient Production Units

Peter Bogetoft
Department of Economics
Royal Agricultural University (KVL)
26 Rolighedsvej, DK-1958 Frederiksberg C
Denmark
e-mail: pb@kvl.dk

Rolf Färe
Department of Economics
Oregon State University
303 Ballard Extension Hall, Corvallis,
97331 Oregon, USA
e-mail: Rolf.Fare@orst.edu

March 1999

Abstract

We discuss how to measure allocative efficiency without presuming technical efficiency. This is relevant when it is easier to introduce reallocations than improvements of technical efficiency. We compare the approach to the traditional one of assuming technical efficiency before measuring allocative efficiency. In particular, we develop necessary and sufficient conditions on the technology to ensure consistent measures and we give dual organizational interpretations of the approaches.

Keywords: Technical Efficiency, Allocative Efficiency, Homethetic Technologies, Slack Consumption.

1 Introduction

In the production economic literature, the overall inefficiency of a production plan is usually decomposed into technical and allocative efficiency in the following manner: First, the plan is projected onto the efficient frontier and some measure of the gains from moving from the original to the projected plan is called the technical efficiency. Next, it is determined which plan is optimal with the given prices and the gains from moving from the projected to the optimal plan is called allocative efficiency. One can say that allocative inefficiency is treated as the residual when evaluating the overall performance - first it is determined what can be explained by technical inefficiency and the rest is called allocative inefficiency. The usual justifications for this approach is that the rate of substitution between the production factors is only well-defined at the frontier.

We believe however that it is relevant to consider the other order of decomposition - i.e. to first evaluate the allocative efficiency and next the technical efficiency. In particular it is relevant to examine how such a decomposition could be undertaken and when the two approaches lead to similar results.

First, from a conceptual point of view, we suggest that the interpretations associated with the notions of technical and allocative efficiency lack a theoretical basis if the size of the effects are dependent of the order of decomposition. In a hierarchical organization composed of several production units, for example, technical inefficiency is interpreted as what could be gained by intra-unit changes and allocative efficiency is interpreted as what can be gained by inter-unit changes. Yet if the latter presumes the former, the two concepts are dependent and not well-defined on their own. In particular, if we change our measure of technical efficiency, say because certain reductions in the resource use may be more natural to consider given to the power structure in the organization, we would also change the allocative efficiency. This means that we can substitute between what we assign to technical and allocative problems which in turn make the interpretations unclear. From a theoretical perspective therefore it is interesting to determine under which technological assumptions the decomposition is unique.

Secondly, from an organizational point of view, we suggest that the two orders of calculations are both equally natural. We show that the traditional decomposition can be thought of in terms of an organization "eating input slacks" while the reversed decomposition we suggest corresponds to the organization "consuming output slack".

Thirdly, from a managerial point of view, it is relevant to determine the allocative efficiency without presuming technical efficiency since it may be relatively easy to reallocate resources within a hierarchy or a market and relatively hard to actually change the production procedures (including the culture, power configuration, incentive structure etc.) used in the individual production units. We therefore need ways to estimate the potential gains from the former without presuming that the latter has already been accomplished. Actually, since most economic productivity studies are relatively aggregate and treat the production units more or less like black boxes, it seems more appropriate in such studies to discuss what can be accomplished by reallocating resources between the units than to examine what can be accomplished by changes inside the units. It seems that latter is the area of organizational specialist while the former is the area of economist in general.

This paper is organized as follows. In Section 2, we formalize some key concepts about the technology. In Section 3, we discuss a general approach to the estimation of allocative efficiency without presuming technical efficiency. Some theoretical results on the consistency of the approach under different technological assumptions are provided in Section 4. Organizational interpretations are emphasized in Section 5, and final remarks are given in Section 6.

2 Technology

We consider the case where a production units has used inputs $x \in \mathbf{R}_+^p$ to produce outputs $y \in \mathbf{R}_+^q$. The *technology* T is given by

$$T = \{(x, y) \in \mathbf{R}_+^p \times \mathbf{R}_+^q | \ x \text{ can produce } y\}$$

and assumed to satisfy the following standard assumptions, cf. e.g. Färe and Primont(1995):

- T is closed
- Inputs and outputs are freely disposable: $(x, y) \in T, x' \geq x, y' \leq y \Rightarrow (x', y') \in T$
- T is convex
- $\{y | (x, y) \in T\}$ is bounded for each $x \in \mathbf{R}_+^p$.

Classical measures of the distance between the production plan (x, y) and the frontier of T is given by Shephard's input and output *distance functions*

$$\begin{aligned} D_i(y,x) &= \sup_{\lambda}\{\lambda > 0 | (x/\lambda, y) \in T\} \\ D_o(x,y) &= \inf_{\theta}\{\theta > 0 | (x, y/\theta) \in T\} \end{aligned}$$

which measures the largest radial contraction of the input vector and the largest radial expansion of the output vector that are feasible in T, cf. Shephard(1953,70). Note that each of these functions give a complete characterization of the technology T since

$$\begin{aligned} D_i(y,x) \geq 1 \Leftrightarrow (x,y) \in T \\ D_o(x,y) \leq 1 \Leftrightarrow (x,y) \in T \end{aligned}$$

We are going to stage most of our discussions in the input space. We note however that a parallel discussion is possible in the output space or even in the full input-output space.

The *input requirement set* $L(y)$ is defined as the set of inputs that can produce output y,

$$L(y) = \{x \in \mathbf{R}^p_+ | \ (x,y) \in T\}$$

and the input *isoquant* $Isoq\ L(y)$ is the sub-set hereof where it is not possible to save on all inputs

$$Isoq\ L(y) = \{x \in L(y) | \varepsilon < 1 \Rightarrow \varepsilon x \notin L(y)\}$$

The *cost function* $C(y, \omega)$ is defined as the minimal cost of producing y when input prices are $\omega \in \mathbf{R}^p_+$

$$C(y,\omega) = \min_x \{\omega \cdot x | x \in L(y)\}$$

Note that by the assumed convexity, there is a dual relationship between the costs and the input distance function in the sense that

$$\begin{aligned} C(y,\omega) &= \min_x \left\{ \frac{\omega x}{D_i(y,x)} \right\} \\ D_i(y,x) &= \inf_\omega \left\{ \frac{\omega x}{C(y,\omega)} \right\} \end{aligned}$$

see Färe and Primont(1995,p.48).

Two properties of the technology will ease the development of alternative decompositions. We say that the technology T is *input-homothetic* if the input sets for different output levels are "parallel", i.e.

$$L(y) = H(y) \cdot \; L(1)$$

Equivalently this means that the cost function can be written as

$$C(y, \omega) = H(y) \cdot \bar{C}(\omega)$$

or that the input distance function equals

$$D_i(y, x) = \bar{D}(x)/H(y)$$

Similarly, we say that the technology T is *input ray-homothetic* if the input sets for different output vectors with the same direction are "parallel"

$$L(y) = \frac{G(y)}{G(y/||y||)} \cdot \; L(\frac{y}{||y||})$$

where $|| \cdot ||$ denote the norm in $\mathbf{R}^q_+$. Equivalently this means that the cost function can be written as

$$C(y, \omega) = \frac{G(y)}{G(y/||y||)} \cdot C(\frac{y}{||y||}, \omega)$$

or that the distance function equals

$$D_i(y, x) = \frac{G(y/||y||)}{G(y)} \cdot D_i(\frac{y}{||y||}, x)$$

3 Efficiency

In the following, we evaluate the production plan (x, y) relative to the technology T and input prices ω. The efficiency indices therefore depends on (x, y), T and ω. To simplify the exposition, however, we do not show this dependence in our notation.

The *overall efficiency OE* of production plan (x, y) in the technology T when input prices are ω is defined as the minimal costs of producing y relative to the actual costs ωx, i.e.

$$OE = \frac{C(y, \omega)}{\omega x}$$

The *standard Farrell* approach views the overall (in)efficiency as originating from two sources, viz.

- technical (in)efficiency, which corresponds to a proportional reduction: $x \to x/D_i(y,x)$ and
- allocative (in)efficiency, which corresponds to the an adjustment to the cost minimal input combination: $x/D_i(y,x) \to x^*(y,\omega) \in \arg\min_{x' \in L(y)} \omega x'$

This leads to the a decomposition of the overall efficiency into *allocative efficiency AE* and *technical efficiency TE* as follows

$$\begin{aligned} OE &= \frac{C(y,\omega)}{\omega x} \\ &= \frac{C(y,\omega)}{\omega\left(\frac{x}{D_i(y,x)}\right)} \cdot \frac{1}{D_i(y,x)} \\ &= AE \cdot TE \end{aligned}$$

We propose here to supplement this standard approach by what we will call a *reverse Farrell approach.* In this we first correct for allocative efficiency and next for technical efficiency:

- allocative (in)efficiency, which corresponds to the an adjustment of the input to a cost minimal input combination: $x \to x^*(\widehat{y},\omega) \in \arg\min_{x' \in L(\widehat{y})} \omega x'$ where $\widehat{y}$ is a reference output vector such that $x \in Isoq\ L(\widehat{y})$, and
- technical (in)efficiency, which corresponds to a proportional reduction: $x^*(\widehat{y},\omega) \to x^*(\widehat{y},\omega)/D_i(y, x^*(\widehat{y},\omega))$

Taking this perspective, we get the following *reversed allocative efficiency* AE^* and *reversed technical efficiency* TE^* measures

$$\begin{aligned} AE^*(\widehat{y}) &= \frac{C(\widehat{y},\omega)}{\omega x} \\ TE^*(\widehat{y}) &= \frac{1}{D_i(y, x^*(\widehat{y},\omega))} \end{aligned}$$

Both approaches, the standard Farrell and the reversed Farrell decompositions, are illustrated in Figure 1 below.

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

FIGURE 1 ABOUT HERE

Figure 1: Two decompositions into allocative and technical efficiency

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

In our new decomposition, there is some freedom in the choice of the output vector at which to evaluate the allocative efficiency. So far we have only assumed that the reference vector $\widehat{y}$ is such that $x \in Isoq\ L(\widehat{y})$. We now discuss this and some alternative assumptions. In general, we denote by $\widehat{Y}$ a non-empty set of reasonable choices.

One possibility is to make no assumptions about $\widehat{y}$ except the general one that

$$x \in IsoqL(\widehat{y}) \tag{A1}$$

There is no obvious "best" way to choose $\widehat{y}$. Indeed, this is precisely what makes the calculation of allocative efficiency without having technical efficiency difficult. Using a sensitivity or lack of information (ignorance) argument one can therefore require the reversed allocative and technical efficiency concepts to give reasonable results for all $\widehat{y}$ satisfying A1.We denote this set $\widehat{Y}(A1)$, i.e. $\widehat{Y}(A1) = \{\widehat{y} \in \mathbf{R}^q_+ | x \in IsoqL(\widehat{y})\}$, and use similar notation for other such sets.

Another possibility is to assume that we can use any $\widehat{y}$ such that

$$x \in IsoqL(\widehat{y}) \text{ and } \widehat{y} \geq y \tag{A2}$$

in the reversed evaluations. One way to rationalize this assumption is in terms of the economics of organizations. One can assume that the production unit has produced $\widehat{y}$ but that it have consumed $\widehat{y} - y$ as slack, either directly by actually producing $\widehat{y}$ and eating $\widehat{y} - y$ or indirectly by supplying less effort than what is needed to actually produce $\widehat{y}$, cf. also Section 5.

Of course, (A2) could be refined in a number of ways by making more precise conjectures about which of the production plans $\widehat{y} \geq y$ the productive unit has actually chosen. A natural refinement in this case is to assume that the production unit has consumed outputs proportionally. This leads to the Farrell selection $\widehat{y} = \widehat{\lambda} y$, and the assumption that $\widehat{y} = \widehat{\lambda} y \in \mathbf{R}^q_+$ such that

$$x \in IsoqL(\widehat{\lambda} y) \tag{A3}$$

Note that whichever assumption we impose on the choice of $\widehat{y}$, we want $\widehat{Y}$ to be non-empty. This is an existence condition.

4 Consistency Properties

In this section, we show when the choice of comparison basis $\widehat{y}$ according to the principles (A1), (A2) and (A3) lead to reversed allocative and technical efficiencies with desirable properties. We commence by discussing some desirable consistency properties.

One requirement that is reasonable in general is the *weak consistency requirement*

$$OE = AE^*(\widehat{y}) \cdot TE^*(\widehat{y}) \qquad \forall \widehat{y} \in \widehat{Y}$$

This requirement says that we want our measures of technical and allocative efficiency to give a (multiplicative) decomposition of the overall efficiency. We note that this is a trivial requirement in the traditional Farrell approach since there allocative efficiency is essentially defined as a residual. It is not trivial here, however, because neither of the components are residuals.

The weak consistency requirement is a guarantee that we do not get too peculiar results in the sense that the combined effect exceeds or fall short of the overall efficiency OE. As such it is a mild requirement.

In particular, it does not restrain the individual effects, only their product. Depending on the technology and the choice of the reference output vector $\widehat{y}$, there may still be lots of room for what to assign to the allocative and technical efficiency components - i.e. to substitute between the two effects. Also, the order of decomposition may still be important. As argued in the introduction, such possibilities to vary the results make the interpretation and application of the technical and allocative efficiency concepts ambiguous. A natural refinement of the weak consistency requirement is therefore to require that the traditional and the reversed Farrell decomposition lead to the same measures of allocative and technical efficiency. We call this the *strong consistency requirement*

$$AE = AE^*(\widehat{y}) \quad \text{and} \quad TE = TE^*(\widehat{y}) \qquad \forall \widehat{y} \in \widehat{Y}$$

With strong consistency, the measures of technical and allocative efficiency neither depend on the order of decomposition nor involve any arbitrariness in the substitution between them. This allows the most clear and unambiguous interpretations of the concepts.

Although the strong consistency requirement is in general stronger than the weak consistency requirement, we shall show below that they do in fact impose the same regularities on the technology for reasonable choices of the $\widehat{Y}$ set.

Initially, we note that one part of the weak consistency requirement is actually trivially. The saving potentials calculated by the reversed measures can never exceed the actual saving potentials. We record this as a lemma.

Lemma 1 *We have*

$$OE \leq AE^*(\widehat{y}) \cdot TE^*(\widehat{y}) \qquad \forall \widehat{y} \in \widehat{Y}(A1)$$

Proof: As above, let $x(y, \omega)$ be a cost minimal input vector capable of producing y when input prices are ω. We then have

$$\begin{aligned} AE^*(\widehat{y}) \cdot TE^*(\widehat{y}) &= \frac{C(\widehat{y}, \omega)}{\omega x} \cdot \frac{1}{D_i(y, x^*(\widehat{y}, \omega))} \\ &= \frac{\omega x^*(\widehat{y}, \omega)/D_i(y, x^*(\widehat{y}, \omega))}{\omega x} \\ &\geq \frac{\omega x^*(y, \omega)}{\omega x} \\ &= OE \end{aligned}$$

because $x^*(\widehat{y}, \omega)/D_i(y, x^*(\widehat{y}, \omega))$ may not be a cost minimizer for y. $\square$

The proof of Lemma 1 suggest that in order for the reversed Farrell measures to be weakly consistent, we basically need the cost minimizer for y and $\widehat{y}$ to be proportional. Depending on the structure of $\widehat{Y}$, this means that we need some type of homothetic technology to get consistent measures. We formalize this in the next propositions.

Proposition 2 *With arbitrary choice of reference output,* $\widehat{y} \in \widehat{Y}(A1)$*, the following properties are equivalent*

1. *Weak consistency*

   $\forall (x, y, \omega) \in T \times \mathbf{R}^p_+ : OE = AE^*(\widehat{y}) \cdot TE^*(\widehat{y}) \quad \forall \widehat{y} \in \widehat{Y}(A1)$

2. *Strong consistency*

   $\forall (x, y, \omega) \in T \times \mathbf{R}^p_+ : AE = AE^*(\widehat{y})$ *and* $TE = TE^*(\widehat{y}) \quad \forall \widehat{y} \in \widehat{Y}(A1)$

3. *T is input homothetic*

Proof: First, assume that the weak consistency requirement holds for an arbitrary $\widehat{y} \in \mathbf{R}^q_+$ such that $x \in IsoqL(\widehat{y})$. From the proof of Lemma 1, we therefore have

$$x^*(y, \omega) = x^*(\widehat{y}, \omega)/D_i(y, x^*(\widehat{y}, \omega))$$

(If there are multiple cost minimizers, we can at least choose $x^*(y, \omega)$ such that this holds). This shows that the expansion path is a ray when we move between any y and $\widehat{y}$ for which there exist an x such that $(x, y) \in T$ and $x \in Isoq(\widehat{y})$. In turn, this implies that T is input homothetic: Let y^1 and y^2 be arbitrary feasible output vectors. Then, by the disposability of inputs, there exist an x such that $(x, y^1) \in T$ and $x \in Isoq(y^2)$ or $(x, y^2) \in T$ and $x \in Isoq(y^1)$. Either way, we get that $x^*(y^1, \omega)$ and $x^*(y^1, \omega)$ are proportional and therefore that T is input homothetic.

Next assume that T is input homothetic. Then the evaluations are also strongly consistent since

$$\begin{aligned}
AE^*(\widehat{y}) &= \frac{C(\widehat{y}, \omega)}{\omega x} \\
&= \frac{C(\widehat{y}, \omega) D_i(\widehat{y}, x)}{\omega x} \\
&= \frac{H(\widehat{y}) \bar{C}(\omega) \bar{D}(x)}{\omega x H(\widehat{y})} \\
&= \frac{H(y) \bar{C}(\omega) \bar{D}(x)}{\omega x H(y)} \\
&= \frac{C(y, \omega)}{\omega x / D_i(y, x)} \\
&= AE
\end{aligned}$$

and

$$\begin{aligned}
TE^*(\widehat{y}) &= \frac{1}{D_i(y, x^*(\widehat{y}, \omega))} \\
&= \frac{1}{D_i(y, x)} \\
&= TE
\end{aligned}$$

since $x^*(\widehat{y}, \omega)$ and x both belong to $L(\widehat{y})$.

Lastly, given strong consistency, weak consistency is trivial since the traditional Farrell measures gives a decomposition of OE:

$$OE = AE \cdot TE$$

$$= \quad AE^*(\widehat{y}) \cdot TE^*(\widehat{y})$$

We have hereby proved (1) $\Rightarrow$ (3) $\Rightarrow$ (2) $\Rightarrow$ (1) and therefore the desired equivalences. $\square$

According to Proposition 1, we need an input homothetic technology to ensure the consistency of the new decomposition for arbitrary choice of the output level $\widehat{y}$ (with $x \in IsoqL(\widehat{y})$).

¿From an organizational point of view, it is easier to justify A2 than A1. It is interesting to note therefore that restricting the choice of output reference to $\widehat{y} \in \widehat{Y}(A2)$ does not affect the conclusions from Proposition 1, including the need for a homothetic technology to ensure that the reversed evaluations are consistent. We record this as a corollary.

Corollary 3 *If we restrict the reference output to exceed the actual output,* $\widehat{y} \in \widehat{Y}(A2)$ *, Proposition 1 is still valid.*

Proof: As in Proposition 1, we may show that (1) $\Rightarrow$ (3) $\Rightarrow$ (2) $\Rightarrow$ (1). The proof of the last three implications from the proof of Proposition 1 is still valid. To show the first implication, assume that the weak consistency requirement holds for an arbitrary $\widehat{y} \in \widehat{Y}(A2)$, i.e. an arbitrary $\widehat{y} \in \mathbf{R}^q_+$ such that $x \in IsoqL(\widehat{y})$ and $\widehat{y} \geq y$. From the proof of Lemma 1, we therefore have

$$x^*(y, \omega) = x^*(\widehat{y}, \omega)/D_i(y, x^*(\widehat{y}, \omega))$$

(If there are multiple cost minimizers, we can choose $x^*(y, \omega)$ such that this holds). This shows that the expansion path is a ray when we move between any y and $\widehat{y} \geq y$ for which there exist an x such that $(x, y) \in T$ and $x \in Isoq(\widehat{y})$. In turn, this implies that T is input homothetic: Let namely y^1 and y^2 be arbitrary feasible output vectors and let $x^1 \in Isoq\ L(y^1)$ and $x^2 \in Isoq\ L(y^2)$. Then, by the disposability of inputs, both x^1 and x^2 can produce $y = \min\{\ y^1,\ y^2\}$ where $\min\{.\}$ is the coordinate-wise minimum. Letting $x = x^1$ and $\widehat{y} = y^1$ we get

$$x^*(y, \omega) = x^*(y^1, \omega)/D_i(y, x^*(y^1, \omega))$$

and letting $x = x^2$ and $\widehat{y} = y^2$ we get

$$x^*(y, \omega) = x^*(y^2, \omega)/D_i(y, x^*(y^2, \omega))$$

Therefore the cost minimizers of arbitrary output vectors y^1 and y^2 are proportional

$$x^*(y^1,\omega) = \frac{D_i(y, x^*(y^1,\omega))}{D_i(y, x^*(y^2,\omega))} \cdot x^*(y^2,\omega)$$

which shows that T is input homothetic as desired. $\square$

If we refine the choice of reference vector into $\widehat{y} \in \widehat{Y}(A3)$, i.e. we assume that the reference vector is found by proportional expansion of the actual y, we get the following necessary and sufficient conditions for consistency.

Proposition 4 *If we restrict the reference output to be proportional to the observed output, $\widehat{y} \in \widehat{Y}(A3)$, the following properties are equivalent*

1. *Weak consistency*

   $\forall(x,y,\omega) \in T \times \mathbf{R}^p_+ : OE = AE^*(\widehat{y}) \cdot TE^*(\widehat{y}) \quad \forall \widehat{y} \in \widehat{Y}(A3)$

2. *Strong consistency*

   $\forall(x,y,\omega) \in T \times \mathbf{R}^p_+ : AE = AE^*(\widehat{y}) \text{ and } TE = TE^*(\widehat{y}) \quad \forall \widehat{y} \in \widehat{Y}(A3)$

3. *T is input ray-homothetic*

Proof: First, assume that the weak consistency requirement holds for (any) $\widehat{y} = \widehat{\lambda y} \in \mathbf{R}^q_+$ such that $x \in IsoqL(\widehat{\lambda} y)$. From the proof of Lemma 1, we therefore have

$$x^*(y,\omega) = x^*(\widehat{\lambda} y,\omega)/D_i(y, x^*(\widehat{\lambda} y,\omega))$$

(If there are multiple cost minimizers, we can choose $x^*(y,\omega)$ such that this holds). This shows that the expansion path is a ray when we move between any y and $\widehat{\lambda} y$ for which there exist an x such that $(x,y) \in T$ and $x \in Isoq(\widehat{\lambda} y)$. In turn, this implies that T is input ray-homothetic.

Next assume that T is input ray-homothetic. Then the evaluations are also strongly consistent since

$$\begin{aligned} AE^*(\widehat{y}) &= \frac{C(\widehat{\lambda} y,\omega)}{\omega x} \\ &= \frac{\frac{G(\widehat{\lambda} y)}{G(\widehat{\lambda} y/||\widehat{\lambda} y||)} C(\frac{\widehat{\lambda} y}{||\widehat{\lambda} y||},\omega)}{\omega x} \end{aligned}$$

$$
\begin{aligned}
&= \frac{\frac{G(\widehat{\lambda}y)}{G(y/||y||)}C(\frac{y}{||y||},\omega)}{\omega x} \\
&= \frac{\frac{G(y)}{G(y/||y||)}C(\frac{y}{||y||},\omega)}{\omega x\frac{G(y)}{G(\widehat{\lambda}y)}} \\
&= \frac{C(y,\omega)}{\omega x\frac{D_i(\widehat{\lambda}y,x)}{D_i(y,x)}} \\
&= \frac{C(y,\omega)}{\omega x/D_i(y,x)} \\
&= AE
\end{aligned}
$$

and

$$
\begin{aligned}
TE^*(\widehat{y}) &= \frac{1}{D_i(y,x^*(\widehat{\lambda}y,\omega))} \\
&= \\
&= \frac{1}{D_i(y,x)} \\
&= TE
\end{aligned}
$$

since $x^*(\widehat{\lambda}y,\omega)$ and x both belong to $Isoq(\widehat{\lambda}y)$ and therefore has the same distance to $Isoq(y)$.

Lastly, given strong consistency, weak consistency is trivial since the traditional Farrell measures gives a decomposition of OE:

$$
\begin{aligned}
OE &= AE \cdot TE \\
&= AE^*(\widehat{y}) \cdot TE^*(\widehat{y})
\end{aligned}
$$

We have hereby proved (1) $\Rightarrow$ (3) $\Rightarrow$ (2) $\Rightarrow$ (1) and therefore the desired equivalences. $\square$

According to Proposition 2, we need an input ray-homothetic technology to ensure the consistency of the new decomposition if we choose $\widehat{\lambda}(x)y$ (with $x \in IsoqL(\widehat{\lambda}(x)y)$) as the basis for evaluating the allocative efficiency.

5 Organizational Interpretations

An important rationale for the reversed decomposition is that it may be easier to reallocate resources than to improve technical efficiency. ¿From an organizational point of view, there are two obvious sources of technical inefficiencies.

One is inadequate decision making. An inefficient production unit uses sub-optimal decision making procedures and production practices. This is the typical explanation found in the productivity analysis literature. It implies that performance can be improved by learning the procedures and practices of the efficient units.

The other possible source of inefficiency is related to the conflicts of interest and asymmetric information in a decentralized organization. The inefficient unit may have excessive on the job consumption of resources and it does not motivate sufficient (non-measured) effort to save on the inputs or expand the output. This perspective has been advocated in for example Bogetoft(1994,95,97).

Either way, the two decompositions, the original and the reversed, can be given *dual organizational interpretations.* The traditional Farrell approach involves excessive consumption of inputs or sub-optimal input handling procedures. The reversed Farrell approach involves excessive consumption of output or sub-optimal output handling procedures. We now formalize this.

The traditional Farrell measure of allocative efficiency can be rewritten as

$$\begin{aligned} AE &= \frac{C(y,\omega)}{\omega x / D_i(y,x)} \\ &= \min_{x'}\{\frac{\omega x'}{\omega x / D_i(y,x)} | x' \in L(y)\} \\ &= \min_{x'}\{\frac{\omega x' D_i(y,x)}{\omega x} | x' \in L(y)\} \\ &= \min_{x''}\{\frac{\omega x''}{\omega x} | x'' \in D_i(y,x) L(y)\} \end{aligned}$$

One interpretation of this is that we seek to reduce the cost of producing y but we realize that we actually need inputs in a factor $D_i(y,x)$ in excess of what is truly needed. That is, if we suggest allocating x'' to the production process, we realize that only $x''/D_i(y,x)$ is actually strictly needed for production. The rest

$$x''(1 - \frac{1}{D_i(y,x)})$$

is consumed as slack. Note that in this interpretation, we do not presume that technical inefficiencies have been eliminated. Rather we assume that the technical inefficiencies are on the input side and will continue at the same general level as before.

Consider next the reversed Farrell measure of allocative efficiency. Also, for ease of comparison, let us consider the case where the reference output is chosen as $\widehat{y} = \widehat{\lambda} y$. In this case

$$\begin{aligned} AE^* &= \frac{C(y/D_o(x,y),\omega)}{\omega x} \\ &= \min_{x'}\{\frac{\omega x'}{\omega x}|x' \in L(y/D_o(x,y))\} \end{aligned}$$

One interpretation of this is that we seek to reduce the cost of producing y but we realize that to actually get y, we need to produce $y/D_o(x,y)$. The excess output

$$y(\frac{1}{D_o(x,y)} - 1)$$

is consumed as slack. Again, we do not presume that technical inefficiencies have been eliminated. Rather, we assume that the technical inefficiencies are on the output side and continue at the same general level as before.

We have shown in Proposition 2 that in order for the two perspectives to give consistent values, we need the technology to be input ray-homothetic. In other cases, the reversed measures do not fully capture the saving potentials. Indeed, we know from Lemma 1 that there will typically be some possibility to gain from reallocation after the second step elimination of technical inefficiency. That is, if we do not restrict the technology and if we initially seek to improve performance by a reallocation and next by a (Farrell) improvement of technical efficiency, we may be left with some "second order" allocative inefficiency

$$AAE^*(\widehat{y}) = \frac{\omega x^*(y,\omega)}{\omega x^*(\widehat{y},\omega)/D_i(y,x^*(\widehat{y},\omega))}$$

resulting from the fact that a proportional (Farrell) reduction in $x^*(\widehat{y},\omega)$ may not be the cost minimizer for the production of y, cf. the proof of Lemma 1. This leads to a three factor decomposition of the overall efficiency

$$OE = AE^*(\widehat{y}) \cdot TE^*(\widehat{y}) \cdot AAE^*(\widehat{y}) \qquad \forall \widehat{y} \in \widehat{Y}$$

or

$$OE = \frac{C(\widehat{y},\omega)}{\omega x} \cdot \frac{1}{D_i(y,x^*(\widehat{y},\omega))} \cdot \frac{\omega x^*(y,\omega)}{\omega x^*(\widehat{y},\omega)/D_i(y,x^*(\widehat{y},\omega))} \qquad \forall \widehat{y} \in \widehat{Y}$$

where the first allocative effect, AE^*, does not presume technical efficiency and may therefore be relatively easy to generate while the latter allocative term, AAE^*, does presume technical adjustments. In Figure 1, AAE^* corresponds to OG/OD.

6 Final Remarks

In this paper we have discussed ways of measuring allocative efficiency without presuming technical efficiency. In the traditional Farrell approach, the technical efficiency is evaluate first and the allocative efficiency next. We introduced a reversed Farrell approach, in which the allocative efficiency is evaluated before technical adjustments are introduced.

We identified necessary and sufficient technological regularities for the two approaches to give consistent measures. For natural choices of the output reference, we need input homothetic or at least input ray-homothetic technologies for the two approaches to be equivalent.

We also gave dual organizational interpretations of the approaches. One interpretation is that the traditional Farrell approach presumes consumption of input slack when evaluating the allocative efficiency while the reversed approach presumes consumption of output slack when the allocative gains are evaluated.

Acknowledgments: This paper was written while Peter Bogetoft visited Department of Agricultural and Resource Economics, University of California at Davis. Financial support from The Danish Social Science Foundation is appreciated.

References

[1] Bogetoft, P., *Incentive Efficient Production Frontiers: An Agency Perspective on DEA*, Management Science, 40, pp.959-968, 1994.

[2] Bogetoft, P., *Incentives and Productivity Measurements*, International Journal of Production Economics, 39, pp. 67-81, 1995.

[3] Bogetoft, P., *DEA-Based Yardstick Competition: The Optimality of Best Practice Regulation*, Annals of Operations Research, 73, pp. 277-298, 1997.

[4] Färe, R. and D. Primont, *Multi-Output Production and Duality: Theory and Applications*, Klüwer Academic Publishers, Boston, 1995.

[5] Shephard, R.W., *Cost and Production Functions*, Princeton University Press, Princeton, New Jersey, 1953

[6] Shephard, R.W., *Theory of Cost and Production Functions*, Princeton University Press, Princeton, New Jersey, 1970