Pergamon

Transpn Res.-B, Vol. 32, No. 5, pp. 299–312, 1998

Printed in Great Britain
0191-2615/98 $19.00+0.00

PII: S0191-2615(98)00002-2

RADIAL AND NONRADIAL STATIC EFFICIENCY DECOMPOSITIONS: A FOCUS ON CONGESTION MEASUREMENT

BENOÎT DERVAUX, KRISTIAAN KERSTENS*
LABORES, Université Catholique de Lille, B.P. 109, 60 Boulevard Vauban, F-59016 Lille Cédex, France

and

PHILIPPE VANDEN EECKAUT
CORE, Université Catholique de Louvain, Voie du Roman Pays 34, B-1348 Louvain-la-Neuve, Belgium

(Received 23 June 1997; accepted 18 November 1997)

Abstract—We summarize and extend the existing frontier literature by outlining proper ways of decomposing efficiency using both radial and nonradial efficiency measures when using nonparametric, deterministic frontier models. Our contribution relates to a recent paper of Viton (Transportation Research-B **31**, 23–39, 1997) analysing the efficiency of U.S. multi-mode bus transit systems. In particular, the author evaluates the technical efficiency, congestion and scale properties of these companies. Due to a methodological mistake he could not detect congestion in his sample. Our proposals remedy this problem and focus on the underexplored issue of proper congestion measurement in general and in transport in particular.

Keywords: urban transit, frontiers, technical efficiency, congestion

1. INTRODUCTION

Viton (1997) analyses the technical efficiency and scale properties of U.S. multi-mode bus transit systems using a nonradial efficiency measure on a series of convex nonparametric, deterministic technologies [also known as Data Envelopment Analysis (DEA) models]. Given the importance of unmeasured inefficiency (or 'slacks') in DEA studies of technical and scale efficiencies when using a traditional, radial efficiency measure, we agree with this choice. Several studies have applied nonradial instead of radial technical efficiency measures. For example, the Russell efficiency measure (Färe and Lovell, 1978) selected by Viton (1997) has been applied first in Deller and Nelson (1991) and furthermore in De Borger and Kerstens (1996), Ferrier *et al.* (1994), Piot-Lepettit *et al.* (1997), among others.

But Viton (1997) also evaluates whether there is congestion in these companies, and finds none. The problem is that this finding is entirely a consequence of his choice of efficiency measure. Congestion cannot be evaluated using the nonradial efficiency measure he adopted. This confusion is related to a lack of clarity in the frontier literature on how to use nonradial efficiency measures. Since congestion is an underexplored issue in transport performance studies, we summarize and extend this frontier literature to provide an overview of possible radial and nonradial decompositions of static efficiency.

This note has the following structure. Section 2 defines several subsets of technology and a series of radial and nonradial efficiency measures. Section 3 defines a static efficiency decomposition, accentuating the importance of congestion. It proposes the traditional radial way of measuring the efficiency components. Then, we suggest alternative ways of using nonradial and radial efficiency measures when decomposing overall technical efficiency. Section 4 concludes.

*Author for correspondence. Fax: 00 33 3 20130470; e-mail: kristiaan.kerstens@flse.fupl.assoc.fr

2. TECHNOLOGY AND EFFICIENCY MEASURES

A production technology is defined by the production possibility set containing all feasible input/output vectors: $\mathbf{S} = \{(x, y)|x \text{ can produce } y\}$. The input requirement set associated with this technology denotes all input vectors x capable of producing a given output vector y: $\mathrm{L}(y) = \{x|(x, y) \in \mathbf{S}\}$.

Two subsets denoting production units on the boundary of technology $\mathrm{L}(y)$ are its isoquant:

$$\text{Isoq } \mathrm{L}(y) = \left\{x|x \in \mathrm{L}(y), \lambda x \notin \mathrm{L}(y) \text{for} \lambda \in [0, 1)\right\}$$

and its efficient subset:

$$\text{Eff } \mathrm{L}(y) = \left\{x|x \in \mathrm{L}(y), x' \leqslant x \Rightarrow x' \notin \mathrm{L}(y)\right\}.$$

Obviously, Isoq $\mathrm{L}(y) \supseteq$ Eff $\mathrm{L}(y)$. Both subsets relate respectively to the Farrell (1957) and Koopmans (1951) definition of technical efficiency.

Technical efficiency is traditionally measured in a radial or equiproportional way (Farrell, 1957). The radial input efficiency measure is defined as:

$$\mathrm{DF}_i(x, y) = \min\left\{\lambda|\lambda \geqslant 0, \lambda x \in \mathrm{L}(y)\right\}.$$

$\mathrm{DF}_i(x, y) \in (0, 1]$, with efficient production on Isoq $\mathrm{L}(y)$ represented by unity. Several nonradial alternatives are defined next.

The Russell (see Färe and Lovell, 1978) input technical efficiency measure is specified:

$$\mathrm{R}_i(x, y) = \min\left\{\sum_{i=1}^{m} \lambda_i/m|\lambda_i \in (0, 1], (\lambda_1 x_1, \ldots, \lambda_m x_m) \in \mathrm{L}(y)\right\}.$$

$\mathrm{R}_i(x, y)$ minimizes the arithmetic mean of scalar reductions in each of the inputs. Since λ_i need not equal λ_j(for $i \neq j$), it is nonradial. $\mathrm{R}_i(x, y) = 1$ if the observation is part of Eff $\mathrm{L}(y)$ and $\mathrm{R}_i(x, y)$ always projects inefficient observations on Eff $\mathrm{L}(y)$.

The Zieschang (1984) input measure of technical efficiency is defined as:

$$\mathrm{Z}_i(x, y) = \mathrm{R}_i(x \cdot \mathrm{DF}_i^+(x, y), y) \cdot \mathrm{DF}_i^+(x, y)$$

where

$$\mathrm{DF}_i^+(x, y) = \min\left\{\lambda|\lambda \geqslant 0, \lambda x \in \mathrm{L}^+(y) = \mathrm{L}(y) + R_+^m\right\}$$

An observation is first rescaled radially (by $\mathrm{DF}_i^+(x, y)$) to Isoq $\mathrm{L}^+(y)$, and the resulting input vector is projected to Eff $\mathrm{L}(y)$ according to $\mathrm{R}_i(x, y)$.* Thus, again $\mathrm{Z}_i(x, y) = 1$ if $(x, y) \in$ Eff $\mathrm{L}(y)$ and $\mathrm{Z}_i(x, y)$ projects inefficient observations onto Eff $\mathrm{L}(y)$.

The asymmetric Färe efficiency measure (Färe, 1975; Färe *et al.*, 1983) is defined as:

$$\mathrm{AF}_i(x, y) = \min_{j=1,\ldots,m}\left\{\mathrm{AF}_i^j(x, y)\right\}$$

where

$$\mathrm{AF}_i^j(x, y) = \min\left\{\lambda_j|\lambda_j \geqslant 0, (x_1, \ldots, \lambda_j x_j, \ldots, x_m) \in \mathrm{L}(y)\right\}.$$

Using a two stage minimization process, $\mathrm{AF}_i(x, y)$ takes the minimum over m components $\mathrm{AF}_i^j(x, y)$, where each component $\mathrm{AF}_i^j(x, y)$ seeks to minimize the use of one input while holding

*$\mathrm{L}^+(y)$ is a technology imposing strong input disposability, derived from adding arbitrary vectors of reals to $\mathrm{L}(y)$.

all other inputs fixed. It correctly identifies observations in Eff $L(y)$ as being Koopmans efficient $(AF_i(x, y) = 1 \Longleftrightarrow x \in \text{Eff } L(y))$ But in general, it projects inefficient observations onto the boundary of the technology (not on its two subsets).

Finally, a modified asymmetric Färe efficiency measure is proposed here and defined as follows:

$$\text{MAF}_i(x, y) = \max_{j=1,\ldots,m} \left\{ \text{AF}_i^j(x, y) \right\}$$

$\text{MAF}_i(x, y)$ takes the maximum over m components $\text{AF}_i^j(x, y)$, where the component $\text{AF}_i^j(x, y)$ is defined as before. While its strong similarity with $\text{AF}_i(x, y)$ is clear, $\text{MAF}_i(x, y)$ has weaker properties. It labels observations efficient as soon as they are part of the boundary of technology and it also projects inefficient observations onto this boundary of technology. Other properties are inherited from $\text{AF}_i(x, y)$. It has a straightforward interpretation: it provides an indication of the minimal effort required to join the boundary of a technology. Its definition relative to the boundary of technology turns out to be useful in efficiency decompositions and in particular in detecting the presence of congestion.

The following relation holds: $\text{AF}_i(x, y) \leqslant \text{R}_i(x, y) \leqslant \text{Z}_i(x, y) \leqslant \text{DF}_i(x, y) \leqslant \text{MAF}_i(x, y)$. Properties satisfied by these nonradial efficiency measures as well as other efficiency measures are discussed in detail by Färe *et al.* (1983) and Kerstens and Vanden Eeckaut (1995). $\text{R}_i(x, y)$ and $\text{Z}_i(x, y)$ on the one hand, and $\text{AF}_i(x, y)$ and $\text{MAF}_i(x, y)$ on the other hand, require modification in case of zero input dimensions (see Färe *et al.*, 1983).

3. DECOMPOSING EFFICIENCY: RADIAL VS NONRADIAL

3.1. A static efficiency taxonomy

Färe *et al.* (1983, 1985, 1994) propose the most elaborate static efficiency taxonomy in the literature and define operational measurement procedures to decompose efficiency for convex nonparametric, deterministic technologies (DEA models). Overall efficiency (OE) is defined as a comparison between any production combination and the situation satisfying its behavioural goal. OE is decomposed to provide information on the possible sources of inefficiency. This static decomposition distinguishes private and social goals.

Private goals are defined in terms of the best interest of the producer and relate to the short run. One distinguishes between technical, structural and allocative efficiency and inefficiency. Technical efficiency (TE) in the Farrell (1957) sense is defined as production on Isoq $L(y)$. A producer is technically inefficient if production occurs in the interior of this level set. Allocative efficiency (AE) requires the specification of a behavioural goal (e.g. cost minimization) and is defined by a point on the boundary of the production possibility set satisfying this objective. A producer is allocative inefficient if there is a divergence between, e.g. observed and optimal costs. Structural efficiency (STE) is closely related to TE. A technically efficient producer is structurally efficient if production occurs in the uncongested region of production. It is structurally inefficient if some inputs yield negative marginal products.*

The social goal relates to a possible divergence between the actual and the ideal size of production, requiring production at a point with constant returns to scale. A productive activity is scale efficient (SCE) if its scale of production corresponds to that resulting from a long run competitive equilibrium; it is scale inefficient otherwise.

Overall technical efficiency (OTE) can be introduced as the union of TE, STE, and SCE. This yields: OTE = TE.STE.SCE.

Figure 1 depicts three input requirement sets and their boundaries all producing the same output level. $L(y)^{sd-crs}$ is characterised by constant returns to scale (CRS) and it is uncongested. $L(y)^{sd-vrs}$ postulates variable returns to scale (VRS) and is also congestion free. $L(y)^{wd-vrs}$ assumes VRS, but allows for congestion. Definitions for these technologies are in Färe *et al.* (1983) and also in the Viton (1997) article. The figure also distinguishes between Isoq $L(y)$ and Eff $L(y)$. Eff $L(y)$

*Modelling congestion requires the assumption of weak instead of strong disposability [see Färe *et al.* (1985) for technical details].

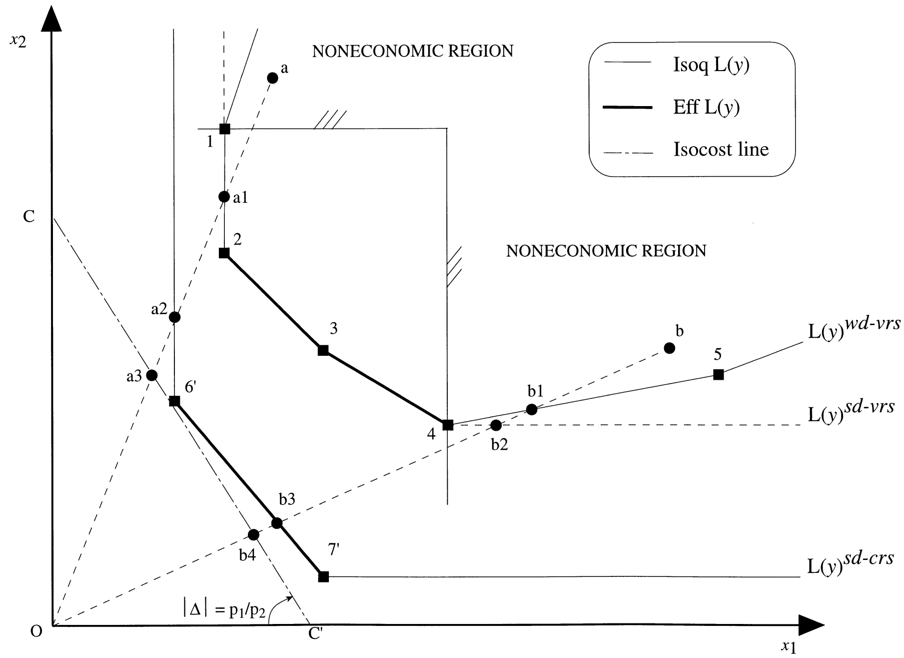

Fig. 1. OE and Subsets of L(y): Radial Decomposition.

consists of the connected line segments 234 and 6′7′ for the VRS respectively CRS technologies. Observe that $L(y)^{sd-vrs}$ and $L(y)^{wd-vrs}$ have an identical Eff L(y).

3.2. The intuition behind efficiency concepts and the importance of congestion

We expand a bit on the intuition behind these static efficiency concepts. Technical efficiency, allocative and scale efficiency seem to be pretty straightforward concepts. Congestion is a less well known phenomenon that most often has been ignored *a priori* by assuming that organizations do not opt for production combinations which yield negative marginal productivity. But this is begging the question.

Congestion in inputs implies that adding more of an input actually decreases total production or requires also more of other inputs to maintain current production levels. This can be easily illustrated on Fig. 1 for point 5. Starting from point 5, increasing input x_1 either leads to a reduced output level or requires more of the second input to remain on the level set. Still otherwise stated, organizations would benefit from simply reducing their congesting input dimensions. In Fig. 1 all observations in the noneconomic region, i.e. that part of technology where at least one input dimension yields a negative marginal product, are subject to congestion.* Well-known examples of congestion are found in agriculture (excessive use of pesticides burns crops) and in transport (traditional transportation networks can only carry a certain level of traffic before their throughput decreases).†

For managerial purposes, it is good to distinguish between short and long term ideals. The order in which the decomposition is defined and measured is indeed dictated by the time perspective of organizational decision making. TE and STE are deemed to be attainable in the short run, since they mainly involve eliminating managerial inefficiencies. SCE and AE are long run goals: they may require scale adjustments respectively changes in the input mix.‡

To which extent is performance measurement an issue in the transport sector? There exists a rather important literature evaluating the performance of urban transit companies in different countries and institutional settings. Part of it was undoubtedly inspired by drastic policy changes

*There is a slight difference in distinguishing economic and noneconomic regions of production between so-called neoclassical (see e.g. Ferguson, 1969, p. 11) and axiomatic (followed in the text) approaches to production.

†The famous speed-flow relationship in traffic theory (see Haight, 1963,p. 72 or May, 1990, p. 284) has recently been interpreted in terms of production theory by Wilson (1991). For road transport, the output speed is produced by a fixed scale of the road and a variable volume (flow) of cars. This economic production model allows for congestion in the output (speed) when the input (volume) is varied.

‡Therefore, the MacDonald (1996) remark on the impact of the order of computation on the decomposition results ignores the economic rationale underlying the originally proposed order.

in several countries following the growing awareness that regulatory failures may impede the correction of existing market failures.

Frontier studies in urban transit include Chu *et al.* (1992), Fazioli *et al.* (1993), Sakano and Obeng (1995), Viton (1986, 1992), among others.* But so far none of these contributions evaluated the presence of congestion. The question of congestion in urban transit companies is an interesting one.

Viton (1997) decomposes OTE of a sample of 217 U.S. motor-bus transit operators providing conventional and/or demand-responsive services in 1990. The mean value of technical efficiency is 0.961 when measured in the input orientation. This means that on average the industry could do with 4% less inputs. The article does not report SCE results, but it provides qualitative information about the number of operators subject to increasing, constant, and decreasing returns to scale. It also fails to detect any congestion (STE). In another decomposition study, Kerstens (1996) found limited levels of congestion in an urban transit sample of French bus companies. The main causes of impaired performance were technical and scale inefficiencies. It also reports qualitative information regarding returns to scale.

However, this single study should not be taken to imply that congestion is a negligible source of poor performance in general or in the transport sector in particular. In the empirical literature we are aware of at least four studies where congestion was, on average, even the most important source of poor performance. More specifically, these studies are Byrnes and Färe (1987) and Byrnes *et al.* (1988) analysing U.S. surface coal mines, Çakmak and Zaim (1992) on Turkish agriculture, and Färe *et al.* (1989) on U.S. electric utilities. So there is definitely scope for more studies focusing on congestion in organizations in the transport sector.

There are at least two possibilities to proceed with measuring these static efficiency concepts: radially and nonradially. We first specify the radial decomposition, and then discuss the possibilities for using nonradial efficiency measures in this static decomposition.

3.3. The radial decomposition of overall efficiency

The traditional radial decomposition uses $\mathrm{DF}_i(x, y)$ relative to different frontiers. By taking appropriate ratios the complete efficiency taxonomy is multiplicatively decomposed. This traditional way of radial input decomposition is illustrated on Fig. 1 for observation b in the interior of $\mathrm{L}(y)^{wd-vrs}$. First, TE is represented by the ratio of distances Ob/Ob1 and is projected to Isoq $\mathrm{L}(y)^{wd-vrs}$. If both inputs of point b are reduced according to the scalar Ob/Ob1, then the resulting input vector b1 is TE in the Farrell (1957) sense. Second, STE is measured by the ratio Ob2/Ob1, derived from comparing the radial distance between a unit without congestion b2 on Isoq $\mathrm{L(y)}^{\mathrm{sd-vrs}}$ and an activity with congestion b1 on Isoq $\mathrm{L}(y)^{wd-vrs}$. Third, SCE is defined by the ratio Ob3/Ob2, i.e. by comparing short and long run isoquants: Isoq $\mathrm{L}(y)^{sd-vrs}$ and Isoq $\mathrm{L}(y)^{sd-crs}$. Fourth, AE is captured by the ratio Ob4/Ob3 indicating the cost reduction from reallocating inputs from point b3 to b4. While b4 cannot yield output y on Isoq $\mathrm{L}(y)^{sd-crs}$ the input vector $6'$ is available for the same budget that can produce this output. Finally, OE is defined as the ratio Ob4/Ob and indicates the total cost reduction possible from moving production from b to b4. Thus, OE = TE.STE.SCE.AE. Using the overall technical efficiency definition, we can also write: OE = OTE.AE.

3.4. Nonradial and almost nonradial decompositions of overall technical efficiency

Nonradial efficiency measures have been proposed in the literature for parts of this static decomposition. As stated before, TE can also be evaluated using $\mathrm{R}_i(x, y)$ to obey the Koopmans instead of the Farrell definition. Also SCE and AE can be decomposed in this way (Färe *et al.*, 1985; p. 148, 1994; p. 82–83). But, it is impossible to measure congestion using $\mathrm{R}_i(x, y)$ since it always projects observations on Eff $\mathrm{L}(y)$. Consequently, measuring efficiency relative to strongly or weakly disposable technologies makes no difference, since they have the same Eff $\mathrm{L}(y)$ (see Fig. 1). Therefore, Viton (1997) is unable to detect any congestion in his sample and in fact lumps together STE and TE.†

*An extensive overview of production studies on urban transit, including performance studies, is found in Berechman (1993). Production and cost studies (including frontier studies of efficiency and productivity) on all transport modes, except urban transit, are surveyed in Oum and Waters (1996).

†In his notation: $\mathrm{R}_i^j(\mathrm{VRS, W}) = \mathrm{R}_o^k(\mathrm{VRS, S})$, by definition.

Already Färe *et al.* (1983; p. 187) stressed that their radial measurement of congestion, explained above, essentially is a way to describe a nonradial phenomenon: congestion results from the excessive usage of one or a subset of inputs, and it need not affect all inputs simultaneously.*

We here propose some ways to reconcile the traditional radial measurement with the desire to eliminate any remaining slacks.† To illustrate the problem of slacks or unmeasured inefficiency, we return to Fig. 1. The radial OTE decomposition of observation a would project onto point a2. However, this point is not part of Eff $L(y)^{sd-crs}$, hence an amount of slack equal to the line segment a2–6′. Similarly, the TE component projects onto a1 on Isoq $L(y)^{sd-vrs}$ leaving a slack amount a1–2.

First it is necessary to state, however, that we think it is not appropriate to use nonradial efficiency measures when prices are available and AE can be evaluated.‡ The reason is that nonradial efficiency measures have an 'implicit' cost interpretation. Their implicit prices are based on the input usage of the observation being evaluated.§ It would be strange to employ one set of prices for decomposing OTE and another one to evaluate AE. Therefore, the following discussion is limited to a decomposition of OTE.

In fact, we propose four ways of decomposing OTE: two proposals yield an entirely nonradial decomposition; the other two lead to a mixture of radial and nonradial efficiency measures.

The two ways to arrive at an entirely nonradial decomposition are to use the asymmetric Färe or the modified asymmetric Färe efficiency measures, since both project inefficient observations onto the boundary. When $AF_i(x, y)$ [or $MAF_i(x, y)$] is evaluated relative to $L(y)^{wd-vrs}$, $L(y)^{sd-vrs}$ and $L(y)^{sd-crs}$, then TE is directly obtained from the first comparison while the STE and SCE components can be computed by taking the appropriate ratios. But it should be noted that the nonradial decomposition based on $AF_i(x, y)$ has a limited application. It only detects STE for congested observations in the interior of the technology, not for congested observations that are efficient in the Farrell (1957) sense.‖ An example of the latter is observation 5 on Fig. 1. We include the decomposition based on the existing $AF_i(x, y)$ efficiency measure mainly to contrast it with the decomposition based on the newly defined $MAF_i(x, y)$ efficiency measure.¶

The other two proposals use $DF_i(x, y)$ to measure TE relative to $L(y)^{wd-vrs}$. The third and fourth decomposition use $R_i(x, y)$ respectively $Z_i(x, y)$ for the STE and SCE components. That is, $R_i(x, y)$ [or $Z_i(x, y)$] are used to evaluate an observation relative to $L(y)^{sd-crs}$ and $L(y)^{sd-vrs}$. Then for the STE and SCE components, ratios are defined as before. This implies that the STE component mixes radial and nonradial efficiency measures, while the SCE component is solely based on nonradial efficiency measures. Their use still guarantees a multiplicative decomposition of OTE, while at the same time eliminating any slacks in the STE and SCE component measures.** The latter decompositions mix radial and nonradial efficiency measures and could therefore be labelled almost nonradial.

The question, which of these four decompositions to choose, depends on the purpose of the exercise. If the purpose is to have a consistent nonradial decomposition, then the first two decompositions are in favour. If the purpose of decomposing OTE is to project observations onto Eff $L(y)$ for the SCE and STE components, then the latter two almost nonradial decompositions are preferred. Given the preoccupation in the frontier literature with eliminating slacks, perhaps the latter two decompositions are most useful.

---

*This nonradial nature of congestion has led to refined radial measurement schemes determining subsets of inputs contributing to the phenomenon: see Byrnes *et al.* (1998) and Färe *et al.* (1994).

†These proposals do not exhaust the possibilities. Cooper *et al.* (1996; pp. 17–21), for instance, proposed another way to measure congestion based on slack variables. This type of efficiency measure, however, lacks independence of units of measurement.

‡In the public sector market prices may not be available. Sometimes shadow prices, which may diverge from market prices, can be determined. For instance, unemployment bounds the social cost of labour below by the unemployment insurance benefit, while its upper bound depends on the extent to which public sector labour demand affects total employment or just displaces private labour demand (Marchand *et al.*, 1984).

§As indicated in Zieschang (1984; pp. 391–392) and explicitly shown in Kerstens and Vanden Eeckaut (1995).

‖This limitation is a consequence of $AF_i(x, y)$ obtaining the unit value only for observations being part of Eff $L(y)$.

¶By computing the partial efficiency measures $AF_i^j(x, y)$ on both weakly and strongly disposable technologies and taking ratios, one has a straightforward tool to detect the presence of congestion in any dimension. This can lead to the determination of subsets of inputs contributing to congestion. Actually, from a computational point of view this method is easier than the one mentioned in footnote * (this page).

**Färe *et al.* (1994; p 83) similarly propose to eliminate slacks in a decomposition of OE by factoring the AE component into a pure allocative component and a slack component [see Ray and Kim (1995) for an empirical application].

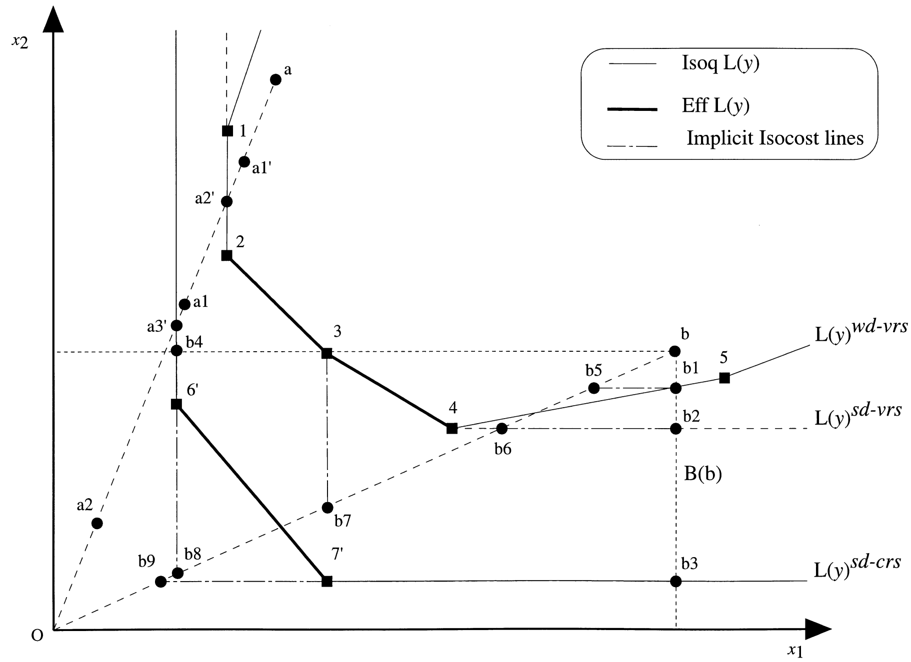

Fig. 2. OTE and Subsets of L(y): Nonradial Decomposition.

A nonradial decomposition based on $\mathrm{AF}_i(x, y)$ and $\mathrm{MAF}_i(x, y)$ is illustrated on Fig. 2. We focus on the OTE decomposition of observation b. Both $\mathrm{AF}_i(x, y)$ and $\mathrm{MAF}_i(x, y)$ are based upon the partial or component efficiency measures $\mathrm{AF}_i^j(x, y)$. We illustrate both efficiency measures for observation b. The components $\mathrm{AF}_i^j(x, y)$ in the first input dimension project observation b onto point 3 and point b4 for the variable respectively constant returns to scale technologies. Note that weakly and strongly disposable technologies share a common projection point. The second input components $\mathrm{AF}_i^j(x, y)$ project observation b onto points b1, b2 and b3 for the weakly disposable, variable returns to scale technology; the strongly disposable, variable returns to scale technology; and the strongly disposable, constant returns to scale technology respectively. Computing now the $\mathrm{AF}_i(x, y)$ and $\mathrm{MAF}_i(x, y)$ efficiency scores relative to the weakly disposable, variable returns to scale technology yields projections onto point 3 respectively point b1. This follows from the maximizing respectively minimizing nature of their objective functions in the second step. For the strongly disposable, variable and constant returns to scale technologies the projection points for $\mathrm{AF}_i(x, y)$ are 3 and b4 and for $\mathrm{MAF}_i(x, y)$ b2 and b3 respectively.

These partial efficiency measures $\mathrm{AF}_i^j(x, y)$ can be geometrically depicted by implicit isocostlines per dimension. Each partial efficiency measure minimizes the implicit cost share of an input dimension relative to the intersection of on the one hand the chosen technology and on the other hand the set of observations doing better than the observation being evaluated (denoted $\mathrm{B}(x^o, y^o) = \{(x, y) | x \leqslant x^o \text{and} y \geqslant y^o\}$). For instance, the implicit isocostlines for observation b in the first and second input dimensions relative to the weakly disposable, variable returns to scale technology are depicted by the line segments 3–b7 respectively b5–b1. The intersection of these implicit isocostlines with the ray through observation b allows a radial reinterpretation of both the efficiency measures $\mathrm{AF}_i(x, y)$ and $\mathrm{MAF}_i(x, y)$. For example, the efficiency measure $\mathrm{AF}_i(x, y)$ relative to the weakly disposable, variable returns to scale technology is the ratio Ob7/Ob. In fact, $\mathrm{AF}_i(x, y)$ takes the minimum of the intersections b7 and b5 of the corresponding isocostlines with the ray.

An almost nonradial decomposition based on $\mathrm{R}_i(x, y)$ is illustrated in detail in Fig. 3. Decomposing again OTE for observation b yields the same TE component Ob/Ob1 as in Fig. 1. But STE and SCE are now both smaller: Ob2/Ob1 respectively Ob3/Ob2. Thus evaluating b with $\mathrm{R}_i(x, y)$ implies implicit prices as indicated by the implicit isocostlines on the figure and leads to a higher estimated amount of congestion and scale inefficiencies.

Obviously, the relative importance of the different sources of inefficiency depends on the choice of efficiency measure. For the traditional radial and the two almost nonradial decompositions, the size of the STE component trivially depends on the efficiency measure in the numerator:

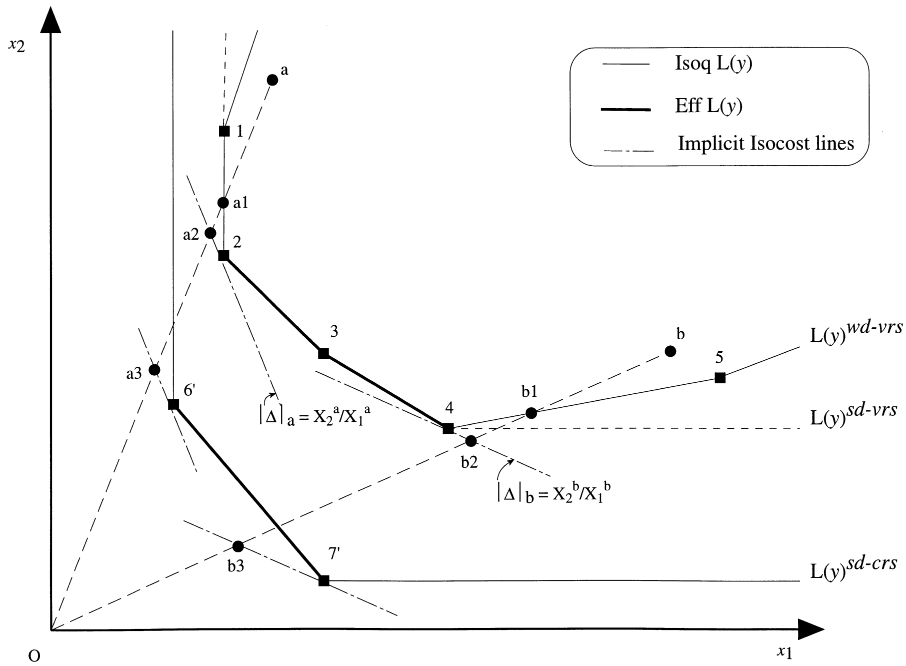

Fig. 3. OTE and Subsets of L(y): Almost Nonradial Decomposition.

$$\frac{\mathrm{R}_i(x, y)^{sd-vrs}}{\mathrm{DF}_i(x, y)^{wd-vrs}} \leqslant \frac{\mathrm{Z}_i(x, y)^{sd-vrs}}{\mathrm{DF}_i(x, y)^{wd-vrs}} \leqslant \frac{\mathrm{DF}_i(x, y)^{sd-vrs}}{\mathrm{DF}_i(x, y)^{wd-vrs}}$$

where superscripts have been added to indicate the technology relative to which observations are being evaluated. Nothing can be said about the relative magnitude of the STE component when ratios of asymmetric Färe or modified asymmetric Färe efficiency measures are taken. No such relation between radial, nonradial, and almost nonradial decompositions can be derived for the SCE component.

4. A NUMERICAL EXAMPLE

Table 1 shows the data for this simple numerical example. These are compatible with observations 1–5 and 6′–7′ on the boundaries of the technologies (represented by squares) on Figs 1–3. Observations a and b are situated in the interior of these technologies (represented by circles). All observations in the Table, except 6 and 7, produce the same level of output. The projection points of observations 6 and 7 on the same output level as the other observations are denoted 6′ and 7′. The original observations 6 and 7 belong to an isoquant with a lower output. All efficiency measures are computed relative to the three technologies mentioned before for the observations 1–5 at output level 5 and observations 6 and 7 with an output of 4.

Table 1. Output and input data

Observations	Output	Input 1	Input 2
1	5	12	34
2	5	12	26
3	5	20	20
4	5	28	14
5	5	48	18
6′	5	8	16
7′	5	20	4
a	5	16	38
b	5	44	20
6	4	6.4	12.8
7	4	16	3.2

Table 2. Radial, nonradial and almost nonradial decomposition results

Observations	Radial decomposition				Almost nonradial decomposition (using $R_i(x, y)$)			
	OTE	TE	SCE	STE	OTE	TE	SCE	STE
1	0.667	1.000	0.667	1.000	0.569	1.000	0.644	0.882
2	0.667	1.000	0.667	1.000	0.641	1.000	0.641	1.000
3	0.600	1.000	0.600	1.000	0.600	1.000	0.600	1.000
4	0.571	1.000	0.571	1.000	0.500	1.000	0.500	1.000
5	0.364	1.000	0.468	0.778	0.319	1.000	0.469	0.681
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
a	0.500	0.750	0.667	1.000	0.461	0.750	0.642	0.956
b	0.375	0.750	0.536	0.933	0.327	0.750	0.490	0.891
	Nonradial decomposition (using $AF_i(x, y)$)				Nonradial decomposition (using $MAF_i(x, y)$)			
1	0.353	0.765	0.462	1.000	0.667	1.000	0.667	1.000
2	0.462	1.000	0.462	1.000	0.667	1.000	0.667	1.000
3	0.200	1.000	0.200	1.000	0.400	1.000	0.400	1.000
4	0.286	1.000	0.286	1.000	0.357	1.000	0.357	1.000
5	0.167	0.472	0.353	1.000	0.222	1.000	0.286	0.778
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
a	0.211	0.605	0.348	1.000	0.500	0.838	0.667	0.895
b	0.182	0.455	0.400	1.000	0.200	0.860	0.286	0.814

Table 2 presents the results for this numerical example obtained for four different decompositions: one radial, one almost nonradial, and finally two nonradial ones. Looking at the figures, it is clear that crucial results to look after in detail are the following. It is obvious that observations a, b and 5 suffer from congestion. Note that the congestion of observation a remains unnoticed by the radial decomposition. Only activities 6 and 7 are scale efficient; all others are scale inefficient. Observations 1 and 5 are technically efficient when using a radial efficiency measure, but are Koopmans inefficient due to slacks.

Turning to the congestion results, the Tables show that observations a, b and 5 are correctly identified as being subject to congestion by the almost nonradial decomposition and by the nonradial decomposition based upon the $MAF_i(x, y)$ efficiency measure. By contrast, the nonradial $AF_i(x, y)$ efficiency measure fails, for this specific example, to reveal any congestion at all. While, as mentioned before, this failure is inevitable for observation 5, it is coincidental for observations a and b. Furthermore, and as can be expected from the earlier derived relation, the ratio of STE diminishes when using an almost nonradial instead of a radial decomposition. Also the use of the $MAF_i(x, y)$ efficiency measure leads to a larger amount of congestion compared to the radial decomposition.

The TE component declines when using the $AF_i(x, y)$ efficiency measure, and increases when applying $MAF_i(x, y)$. This follows from the weak ordering between efficiency measures derived above. By definition, radial and almost nonradial decompositions have an identical TE component. Observations 6 and 7 are correctly identified as the only scale efficient activities in all decompositions. The SCE component happens to decline for all alternatives to the radial decomposition. The OTE component, resulting from multiplying TE, STE and SCE, also tends to decline for all alternative decompositions.

How do the different decompositions deal with the slacks left by radial projections? As stated before, only the almost nonradial decompositions attempt to deal with slacks. The nonradial decompositions aim at consistency in the use of nonradial measurement. We can therefore concentrate on the former. For observation 5, for instance, the amount of congestion (STE) increases under the almost nonradial decomposition, because it incorporates the slacks unmeasured by the radial TE component. The same holds true for observation b. Observation 1 now even appears as being congested, because the slacks relative to the variable and constant returns to scale technologies are now partly counted as congestion and partly attributed to SCE.

The reader can verify the results listed in Table 2 on the figures using a simple ruler, by exploiting the radial reinterpretation of the nonradial efficiency measures outlined previously. For instance, TE for observation b using $\mathrm{MAF}_i(x, y)$ is 0.860. On Fig. 2 this is (approximately) equal to the ratio of distances Ob5/Ob. As another example, using $\mathrm{R}_i(x, y)$ the SCE for observation a is 0.642. This value (approximately) results from computing the ratio Oa3/Oa2 in Fig. 3.

Concluding, the numerical example underscores the potentials of the nonradial and almost nonradial decompositions as alternatives to the traditional, radial decomposition. Suffice to add that details on the algorithms for implementing the radial and nonradial efficiency measures, defined in Section 2, on convex DEA models are provided in Ferrier *et al.* (1994). They are also discussed at great length in the Appendix.

5. CONCLUSIONS

The methodological error in Viton (1997), when exploring the presence of congestion in U.S. multi-mode bus transit systems, has led to a reconsideration of the use of nonradial efficiency measures in static efficiency decompositions. For nonparametric, deterministic frontier models, a number of possibilities for decomposing OTE in both radial, nonradial and almost nonradial ways have been discussed. These alternative decompositions allow in principle to break down OTE into its three components.

Acknowledgements—The authors thank J. Berechman for his helpful comments. The authors bear full responsibility for any remaining errors.

REFERENCES

Berechman, J. (1993) *Public Transit Economics and Deregulation Policy*. Amsterdam, North-Holland.

Byrnes, P. and Färe, R. (1987) Surface mining of coal: efficiency of US interior mines. *Applied Economics* **19**, 1665–1673.

Byrnes, P., Färe, R., Grosskopf, S. and Lovell, C. A. K. (1988) The effect of unions on productivity: U.S. surface mining of coal. *Management Science* **34**, 1037–1053.

Çakmak, E. and Zaim, O. (1992) Technical efficiency of Turkish agriculture. *Middle East Technical University Studies in Development* **19**, 305–316.

Chu, X, Fielding, G. and Lamar, B. (1992) Measuring transit performance using data envelopment analysis. *Transportation Research* **26A**, 223–230.

Cooper, W., Thompson, R. and Thrall, R. (1996) Introduction: extensions and new developments in DEA. *Annals of Operations Research* **66**, 3–45.

De Borger, B. and Kerstens, K. (1996) Radial and nonradial measures of technical efficiency: an empirical illustration for belgian local governments using an FDH reference technology. *Journal of Productivity Analysis* **7**, 41–62.

Deller, S. and Nelson, C. (1991) Measuring the economic efficiency of producing rural road services. *American Journal of Agricultural Economics* **72**, 194–201.

Farrell, M. (1957) The measurement of productive efficiency. *Journal of Royal Statistical Society* **120A**, 253–281.

Färe, R. (1975) Efficiency and the production function. *Zeitschrift fur Nationalökonomie* **35**(3–4), 317–324.

Färe, R., Grosskopf, S. and Lovell, C. A. K. (1983) The structure of technical efficiency. *Scandinavian Journal of Economics* **85**, 181–190.

Färe, R., Grosskopf, S. and Lovell, C. A. K. (1985) *The Measurement of Efficiency of Production*. Kluwer, Boston.

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

Färe, R., Grosskopf, S. and Pasurka, C. (1989) The effect of environmental regulations on the efficiency of electric utilities: 1969 versus 1975. *Applied Economics* **21**, 225–235.

Färe, R. and Lovell, C. A. K. (1978) Measuring the technical efficiency of production. *Journal of Economic Theory* **19**, 150–162.

Färe, R., Lovell, C. A. K. and Zieschang, K. (1983) Measuring the technical efficiency of multiple output production technologies, In *Quantitative Studies on Production and Prices*, eds. W. Eichhorn, K. Neumann and R. Shephard (eds), Physica-Verlag, Würzburg, pp. 159–171.

Fazioli, R., Filippini, M. and Prioni, P. (1993) Cost-structure and efficiency of local public transport: the case of Emilia Romagna Bus Companies. *International Journal of Transport Economics* **20**, 305–324.

Ferguson, C. (1969) *The Neoclassical Theory of Production and Distribution*. Cambridge University Press, Cambridge.

Ferrier, G., Kerstens, K. and Vanden Eeckaut, P. (1994) Radial and nonradial technical efficiency measures on a DEA reference technology: a comparison using US banking data. *Recherches Économiques de Louvain* **60**, 449–479.

Haight, F. (1963) *Mathematical Theories of Traffic Flow*. Academic Press, New York.

Kerstens, K. and Vanden Eeckaut, P. (1995) Technical efficiency measures on DEA and FDH: a reconsideration of the axiomatic literature. Louvain-la-Neuve, UCL (CORE Discussion Paper 9513).

Kerstens, K. (1996) Decomposing overall technical efficiency of french urban transport, Loughborough, Loughborough University (Dept. of Economics: Economics Research Paper Nr. 96/11).

Koopmans, T. (1951) Analysis of production as an efficient combination of activities. In *Activity Analysis of Production and Allocation*, ed. T. Koopmans, pp. 33–97. Yale University Press, New Haven.

Marchand, M., Mintz, J. and Pestieau, P. (1984) Shadow pricing of labour and capital in an economy with unemployed labour. *European Economic Review* **25**, 239–252.

May, A. (1990) *Traffic Flow Fundamentals*, Prentice Hall, Englewood Cliffs.

MacDonald, J. (1996) Note: a problem with the decomposition of technical inefficiency into scale and congestion components. *Management Science* **42**, 473–474.

Oum, T. and Waters, W. (1996) A survey of recent developments in transportation cost function research. *Logistics and Transportation Review* **32**, 423–463.
Piot-Lepettit, I., Vermersch, D. and Weaver, R. (1997) Agriculture's environmental externalities: DEA evidence for french agriculture. *Applied Economics* **29**, 331–338.
Ray, S. and Kim, H. (1995) Cost efficiency in the US steel industry: a nonparametric analysis using data envelopment analysis. *European Journal of Operational Research* **80**, 654–671.
Sakano, R. and Obeng, K. (1995) Re-examination of inefficiencies in urban transit systems: a stochastic frontier approach. *Logistics and Transportation Review* **31**, 377–392.
Viton, P. (1986) The question of efficiency in urban bus transportation. *Journal of Regional Science* **26**, 499–513.
Viton, P. (1992) Consolidations of scale and scope in urban transit. *Regional Science and Urban Economics* **22**, 25–49.
Viton, P. (1997) Technical efficiency in multi-mode bus transit: a production frontier analysis. *Transportation Research* **31B**, 23–39.
Wilson, P. (1991) The highway speed volume relationship revisited. *Journal of Urban Economics* **29**, 202–206.
Zieschang, K. (1984) An extended Farrell efficiency measure. *Journal of Economic Theory* **33**, 387–396.

APPENDIX

This Appendix presents the mathematical programming problems needed to implement the radial, nonradial and almost nonradial decompositions described in the text. The radial decomposition has already been explained in detail in Färe *et al.* (1985), but is repeated here for convenience and for comparative purposes.

Before sequentially describing the decompositions, it is useful to formally define the three technologies $L(y)^{wd-vrs}$, $L(y)^{sd-vrs}$, and $L(y)^{sd-crs}$ referred to in Section 3.1. The reader can also consult Färe *et al.* (1983, 1985, 1994) for details.

The input correspondence $L(y)^{wd-vrs}$ of the non-parametric deterministic production model with weak disposability in inputs and strong disposability in outputs and with variable returns to scale is constructed from observed activities as follows:

$$L(y)^{\mathrm{wd-vrs}} = \Big\{x| \sum_{k=1}^{K} y_{kn} z_k \geqslant y_n, n = 1, \dots, N, \sum_{k=1}^{K} x_{km} z_k = \mu x_m, m = 1, \dots, M, \sum_{k=1}^{K} z_k = 1, \mu \in (0, 1], z_k \geqslant 0, k = 1, \dots, K\Big\}.$$

The technology with strong disposability in both inputs and outputs and variable returns to scale has the following input correspondence $L(y)^{sd-vrs}$:

$$L(y)^{sd-vrs} = \Big\{x| \sum_{k=1}^{K} y_{kn} z_k \geqslant y_n, n = 1, \dots, N, \sum_{k=1} x_{km} z_k \leqslant x_m, m = 1, \dots, M \sum_{k=1}^{K} z_k = 1, z_k \geqslant 0, k = 1, \dots, K\Big\}.$$

Finally, the strongly disposable technology with constant returns to scale has input correspondence $L(y)^{\mathrm{sd-crs}}$:

$$L(y)^{sd-crs} = \Big\{x| \sum_{k=1}^{K} y_{kn} z_k \geqslant y_n, n = 1, \dots, N, \sum_{k=1K} x_{km} z_k \leqslant x_m, m = 1, \dots, M \; z_k \geqslant 0, k = 1, \dots, K\Big\}.$$

The three programming models in each decomposition always measure efficiency relative to these three technologies, but differ from each other in the type of efficiency measure they are using.

Radial decomposition

Radial efficiency in the inputs is computed relative to a variable returns to scale technology with weak disposability in inputs and strong disposability in outputs by solving for each observation (x^o, y^o) the following programming problem (R.1):

$$DF_i(x, y) = \mathrm{Min}_{\lambda,z} \lambda$$

subject to

$$\sum_{k=1}^{K} y_{kn} z_k \geqslant y^o_{kn}, n = 1, \dots, N \qquad (R.1 - 1)$$

$$\sum_{k=1}^{K} x_{km} z_k = \mu\lambda x_{km}^{o}, \ m = 1, \ldots, M \tag{R.1 – 2}$$

$$\sum_{k=1}^{K} z_k = 1 \tag{R.1 – 3}$$

$$\mu \leqslant 1 \tag{R.1 – 4}$$

$$\lambda \geqslant 0, \mu \geqslant 0, z_k \geqslant 0, k = 1, \ldots, K \tag{R.1 – 5}$$

It is non-linear in the parameters, but can be transformed into a linear programming problem as follows. First, divide the constraints (R.1–1) to (R.1–3) by μ. Second, redefine $z_i' = z_i/\mu$. The resulting problem is now linear in all parameters. The only problem is that (R.1–3) now reads:

$$\sum_{k=1}^{K} z_k' = 1/\mu \tag{R.1 – 3′}$$

The right hand side is no longer equal to unity. But as noted by Färe *et al.* (1985; p. 179), it is possible to restrict the right hand side to be equal to unity. This ensures that the observation (x^o, y^o) can be part of the frontier spanned by the observations. The resulting linear programming problem is straightforward to solve.

Radial efficiency in the inputs is calculated relative to a strongly disposable variable returns to scale technology by solving for each observation (x^o, y^o) the following linear programming problem (R.2):

$$DF_i(x, y) = \mathrm{Min}_{\lambda,z}\lambda$$

subject to

$$\sum_{k=1}^{K} y_{kn} z_k \geqslant y_{kn}^{o}, \ n = 1, \ldots, N \tag{R.2 – 1}$$

$$\sum_{k=1}^{K} x_{km} z_k \leqslant \lambda x_{km}^{0}, \ m = 1, \ldots, M \tag{R.2 – 2}$$

$$\sum_{k=1}^{K} z_k = 1 \tag{R.2 – 3}$$

$$\lambda \geqslant 0, z_k \geqslant 0, kz = 1, \ldots, K \tag{R.2 – 4}$$

Radial efficiency in the inputs is computed relative to a strongly disposable constant returns to scale technology by solving for each observation (x^o, y^o) a linear programming problem (R.3) which is identical to (R.2) except that the constraint (R.2-3) is dropped.

Nonradial decompositions

Input efficiency according to $AF_i(x, y)$ or $MAF_i(x, y)$ requires first solving one linear program per component efficiency measure $AF_i^j(x, y)$. Then the minimum respectively the maximum is taken over these component efficiency measures.

$AF_i^j(x, y)$ is computed relative to a variable returns to scale technology with weak disposability in inputs and strong disposability in outputs by solving for each observation (x^o, y^o) the programming problem (N.1):

$$AF_i^j(x, y) = \mathrm{Min}_{\lambda,z}\lambda_j$$

subject to

$$\sum_{k=1}^{K} y_{kn} z_k \geqslant y_{kn}^{o}, \ n = 1, \ldots, N \tag{N.1 – 1}$$

$$\sum_{k=1}^{K} x_{km} z_k = \mu\lambda_j x_{km}^{o}, \ m = j \tag{N.1 – 2}$$

$$\sum_{k=1}^{K} x_{km} z_k = \mu x^o_{km}, m \neq j \qquad (N.1-3)$$

$$\sum_{k=1}^{K} z_k = 1, m = 1, \ldots, M \qquad (N.1-4)$$

$$\mu \leqslant 1 \qquad (N.1-5)$$

$$\lambda_i \geqslant 0, \mu \geqslant 0, z_k \geqslant 0, k = 1, \ldots, K \qquad (N.1-6)$$

This programming problem can, in contrast to problem (R.1), not be transformed into a linear program.

Efficiency in the inputs according to $AF^j_i(x, y)$ is calculated relative to a strongly disposable variable returns to scale technology by solving for each observation (x^o, y^o) the following linear programming problem (N.2):

$$AF^j_i(x, y) = \mathrm{Min}_{\lambda,z} \lambda_j$$

subject to

$$\sum_{k=1}^{K} y_{kn} z_k \geqslant y^o_{kn}, n = 1, \ldots, N \qquad (N.2-1)$$

$$\sum_{k=1}^{K} x_{km} z_k \leqslant \lambda_j x_{km^o,m} = j \qquad (N.2-2)$$

$$\sum_{k=1}^{K} x_{km} z_k \leqslant x^o_{km}, m \neq j \qquad (N.2-3)$$

$$\sum_{k=1}^{K} z_k = 1, m = 1, \ldots, M \qquad (N.2-4)$$

$$\lambda_j \geqslant 0, z_k \geqslant 0, k = 1, \ldots, K \qquad (N.2-5)$$

The input efficiency measure $AF^j_i(x, y)$ is computed relative to a strongly disposable constant returns to scale technology by solving for each observation (x^o, y^o) a linear programming problem (N.3) that is identical to (N.2) except that there is no longer a constraint (N.2-4).

Almost nonradial decompositions

Radial efficiency in the inputs is computed relative to a variable returns to scale technology with weak disposability in inputs and strong disposability in outputs by solving for each observation (x^o, y^o) the programming problem (AN.1). This is identical to the programming problem (R.1).

Efficiency in the inputs according to $R_i(x, y)$ is calculated relative to a strongly disposable variable returns to scale technology by solving for each observation (x^o, y^o) the following linear programming problem (AN.2):

$$R_i(x, y) = \mathrm{Min}_{\lambda,z} \sum_{m=1}^{M} \frac{\lambda_m}{M}$$

subject to

$$\sum_{k=1}^{K} y_{kn} z_k \geqslant y^o_{kn}, n = 1, \ldots, N \qquad (AN.2-1)$$

$$\sum_{k=1}^{K} x_{km} z_k \leqslant \lambda_m x^o_m, m = 1, \ldots, M \qquad (AN.2-2)$$

$$\sum_{k=1}^{K} z_k = 1 \qquad (AN.2-3)$$

$$\lambda_{\mathrm{m}} \leqslant 1 \tag{AN.2 – 4}$$

$$\lambda_{\mathrm{m}} \geqslant 0, z_{\mathrm{k}} \geqslant 0, \mathrm{k} = 1, \ldots, \mathrm{K} \tag{AN.2 – 5}$$

Observe that the constraint (AN.2-4) is absolutely necessary.

The Russell input efficiency measure $\mathrm{R}_i(x, y)$ is computed relative to a strongly disposable constant returns to scale technology by solving for each observation (x^o, y^o) a linear programming problem (AN.3) that is identical to (AN.2) except that there is no longer a constraint (AN.2-3).

Similar programming problems for the Zieschang efficiency measure $\mathrm{Z}_i(x, y)$ can be developed. The first problem is again identical to the programming problem (R.1). The second problem boils down to a two step procedure. First, a radial efficiency measure is computed according to problem (R.2), yielding λ^* as optimal value. Second, a Russell efficiency measure is calculated for the modified observation $(\lambda^*.x^o, y^o)$ following problem (AN.2). The Zieschang efficiency measure is simply the product of the optimal efficiency measures obtained in these two steps (see also Ferrier *et al.*, 1994). The third programming problem can again be derived from the second.