ORIGINAL RESEARCH



Scale characteristics of variable returns-to-scale production technologies with ratio inputs and outputs

Ole Bent Olesen¹ · Niels Christian Petersen¹ · Victor V. Podinovski²

Accepted: 24 June 2022 / Published online: 16 July 2022 © The Author(s) 2022

Abstract

Applications of data envelopment analysis (DEA) often include inputs and outputs represented as percentages, ratios and averages, collectively referred to as ratio measures. It is known that conventional DEA models cannot correctly incorporate such measures. To address this gap, the authors have previously developed new variable and constant returns-to-scale models and computational procedures suitable for the treatment of ratio measures. The focus of this new paper is on the scale characteristics of the variable returns-to-scale production frontiers with ratio inputs and outputs. This includes the notions of the most productive scale size (MPSS), scale and overall efficiency as measures of divergence from MPSS. Additional development concerns alternative notions of returns to scale arising in models with ratio measures. To keep the exposition as general as possible and suitable in different contexts, we allow all scale characteristics to be evaluated with respect to any selected subsets of volume and ratio inputs and outputs, while keeping the remaining measures constant. Overall, this new paper aims at expanding the range of techniques available in applications with ratio measures.

Keywords Data envelopment analysis \cdot Technology \cdot Ratio measures \cdot Scale efficiency \cdot Returns to scale

1 Introduction

Applications of data envelopment analysis (DEA) often involve inputs and outputs represented by ratio data. Such data may reflect the environment in which the decision making

Victor V. Podinovski v.podinovski@lboro.ac.uk

> Ole Bent Olesen ole@sam.sdu.dk

Niels Christian Petersen ncp@sam.sdu.dk

¹ Department of Business Economics, The University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

² School of Business and Economics, Loughborough University, Loughborough LE11 3TU, UK

units (DMUs) operate or describe the quality of inputs and outputs involved in the production process. For example, in the context of assessment of school performance, ratio measures may include the average income per capita in the catchment area of a school, proportion of pupils with special needs and the percentage of students achieving a certain high level in exams.

The two standard DEA models are based on the assumptions of constant and variable returns-to-scale (CRS and VRS). These models were introduced in DEA by Charnes et al. (1978) and Banker et al. (1984) and can be seen as continuing the earlier developments in the economics literature by Afriat (1972), Shephard (1974), Färe and Lovell (1978) and Färe et al. (1983a). It has long been realised that the standard CRS and VRS DEA models are generally not suitable if some data are given in the form of ratios—see, e.g., Dyson et al. (2001), Cooper et al. (2007), and Emrouznejad and Amin (2009). The main reason for this is that production technologies with ratio data do not generally satisfy the assumption of convexity which is incorporated in the VRS and CRS technologies. In the case of CRS, a further problem arises because ratio measures are generally not scalable in the same way as conventional volume inputs and outputs such as costs, labor and physical levels of production or services. As also noted by Pastor et al. (2013), the input and output projections of inefficient DMUs obtained in the CRS model may result in the target values located outside the range of values observed in the empirical data set. This becomes a particular problem in the case of ratio measures, such as percentages, which may have a natural upper bound such as unity or 100%.¹

In order to overcome the highlighted problems with the use of ratio data in DEA, Olesen et al. (2015) introduced new ratio-VRS (R-VRS) and ratio-CRS (R-CRS) production technologies that incorporate both volume and ratio inputs and outputs as *native types of data*, i.e., without any modification. Both technologies are formally derived from the explicitly stated sets of production axioms. In particular, to allow a different treatment of ratio measures compared to volume measures, Olesen et al. (2015) utilize the axiom of selective convexity of Podinovski (2005), instead of the conventional convexity assumption. In the case of R-CRS, the standard axiom of full proportionality (scalability) is replaced by the axiom of selective proportionality, which also distinguishes between the volume and ratio inputs and outputs.

The R-VRS and R-CRS models with ratio data are further explored by Olesen et al. (2017). The latter paper discusses efficiency concepts in models with ratio data, including the new notion of potential ratio efficiency, and computational approaches for their testing. In the most recent paper on this subject, Olesen et al. (2022) explore the geometric structure of the R-VRS technology and the R-CRS technology with fixed ratio inputs and outputs. In particular, they prove that the R-VRS technology is the union of a finite number of specially constructed standard VRS technologies.

In the current paper we consider scale characteristics of the production frontier of the R-VRS technology that have not yet been explored. This includes the notions of the most productive scale size (MPSS), scale and overall efficiency, and the related notions of returns to scale (RTS). All these notions have *conceptually* been defined and discussed in the literature—see, e.g., Frisch (1965) and Färe et al. (1983b, 1985). Methods of their evaluation in the standard VRS technology (see, e.g., Banker 1984; Banker and Thrall 1992; Førsund and Hjalmarsson 2004; Chambers and Färe 2008) and in the whole large classes of polyhedral and convex technologies (Podinovski et al. 2016; Podinovski 2017) have also been developed. However, these methods do not generally apply to the R-VRS technology (because it is not a

¹ The bounded CRS model of Pastor et al. (2013) addresses this problem by incorporating upper bounds on the outputs (and, similarly, lower bounds on the inputs). See also related work on bounded additive DEA models by Cooper et al. (2011) and Pastor et al. (2015).

convex technology), and even the known conceptual characteristics such as scale efficiency may not appear to be uniquely defined. In our paper we address these issues.

Several contributions of our paper are worth highlighting. First, following the approach of Banker (1984), we define and interpret scale efficiency in the R-VRS technology as a measure of divergence from MPSS. A challenging problem here is that the cone and CRS extensions of the R-VRS technology do not coincide. What appears to be a straightforward extension of the traditional approach to the evaluation of scale efficiency turns out to be a conceptual dilemma.

Second, to make the exposition as general as possible, we give the definition of MPSS and scale efficiency with respect to any selected subsets of volume and ratio inputs and outputs, while keeping the remaining measures constant. Conceptually, this is a generalization of the approach of Banker and Morey (1986) who consider technical and scale efficiency with respect to discretionary inputs and outputs only. However, because the R-VRS technology has a more complex structure than the standard VRS technology, the actual extension of the approach of Banker and Morey (1986) to the R-VRS technology is not straightforward.

In the additional development, we explore the notion of RTS. We differentiate between the cases in which the types of RTS are evaluated with respect to volume inputs and outputs only and the general case involving ratio measures. In the former case, we explore the local characterization of RTS with respect to the selected inputs and outputs. In the latter case, we show that the conventional local RTS characterization becomes trivial and uninformative. Instead, we employ the global RTS characterization of production frontiers developed by Podinovski (2004a, b) whose types are indicative of the direction to MPSS.

The methodology developed in our paper extends the well-known approaches to the evaluation of scale efficiency and RTS to the technologies with ratio inputs and outputs. Furthermore, because our approach allows for a selection of inputs and outputs with respect to which we measure scale characteristics, it now becomes possible to explore the relationship between the volume inputs and outputs, while keeping the socio-economic and quality characteristics of the production process (represented by percentages) constant. Alternatively, it is also possible to explore the relationship between the socio-economic factors and quality characteristics of the production process, for the given levels of volume inputs and outputs. We illustrate the usefulness of such approach in an application to secondary schools in England.

We proceed as follows. In Sect. 2, we introduce basic definitions and notation. In Sect. 3, we briefly outline the R-VRS technology and give a clarifying illustrative example.

Section 4 contains the main theoretical results of our paper. We first define the partial cone extension of the R-VRS technology, which is subsequently used to define MPSS and overall and scale efficiency of DMUs with regard to arbitrary subsets of volume and ratio inputs and outputs. We also show that the cone and CRS extensions of the R-VRS technologies are generally different sets and discuss implications of this result.

In Sect. 5, we operationalize the models developed for the assessment of scale characteristics in the R-VRS technology. In Sect. 6, we show that several known technologies and methods of evaluation of scale characteristics are special cases of our newly developed approach. In Sect. 7, we consider local and global RTS characterizations in the R-VRS technology. In Sect. 8, we consider an application in the context of secondary education which illustrates the evaluation of scale characteristics in the R-VRS technology. Section 9 contains concluding remarks.

All mathematical proofs are given in Appendix A. An additional example clarifying the discussion in Sect. 4 is considered in Appendix B. The full data set used in the application is given in Appendix C. The GAMS code used for computations in the application is available online.

2 Preliminaries

Following the notation introduced by Olesen et al. (2015), let $T \subset \mathbb{R}^m_+ \times \mathbb{R}^s_+$ be a production technology with the sets $I = \{1, ..., m\}$ and $O = \{1, ..., s\}$ of nonnegative inputs and outputs, respectively. We denote $I^V \subseteq I$ and $O^V \subseteq O$ the subsets of *volume* inputs and outputs. The complementary subsets $I^R = I \setminus I^V$ and $O^R = O \setminus O^V$ include *ratio* inputs and outputs. We assume that both sets $I = I^V \cup I^R$ and $O = O^V \cup O^R$ are not empty, although any of their subsets I^V, I^R, O^V and O^R may be empty.

DMUs are elements of technology T and are stated in the form

$$(X, Y) = (XV, XR, YV, YR),$$

where $X \in \mathbb{R}^m_+$ and $Y \in \mathbb{R}^s_+$ are the vectors of inputs and outputs, and their subvectors X^V , X^R , Y^V and Y^R correspond to the sets of volume and ratio measures I^V , I^R , O^V and O^R , respectively.

Let (X_j, Y_j) be observed DMUs, where $j \in J = \{1, ..., n\}$. We assume that each observed DMU has at least one strictly positive input and at least one strictly positive output, i.e., $X_j \neq \mathbf{0}$ and $Y_j \neq \mathbf{0}$, for all $j \in J$. (We use bold symbols **0** and **1** to denote the vectors of zeros and ones whose dimensions are clear from the context.)

Denote DMU (X_o, Y_o) the particular DMU whose efficiency or scale efficiency is being considered. This may be any DMU from technology T, including one of the observed DMUs.

Often, ratio inputs and outputs have certain upper bounds, typically either unity or 100%. Following Olesen et al. (2015), we state these bounds in the form

$$X^R \le \bar{X}^R$$
 and $Y^R \le \bar{Y}^R$, (1)

where each component of vectors \bar{X}^R and \bar{Y}^R can be either finite or $+\infty$. (Vector inequalities mean that the specified inequality is true for each component, e.g., $X^R \leq \bar{X}^R$ means that $X_i^R \leq \bar{X}_i^R$, for all $i \in I^R$.) We naturally assume that the vectors of ratio inputs and outputs X_i^R and Y_i^R of all observed DMUs $j \in J$ satisfy the inequalities (1).

3 The R-VRS technology

Olesen et al. (2015) note that the conventional VRS technology of Banker et al. (1984) should generally not be used if some inputs and outputs are ratio measures. The main reason for this is that ratio measures cannot be assumed to satisfy the axiom of convexity which is explicitly required by the definition of the VRS technology. Instead, Olesen et al. (2015) demonstrate that one should exclude the ratio inputs and outputs from the convexity assumption, while keeping the convexity property only for the volume measures.

For a formal definition of the R-VRS technology, Olesen et al. (2015) assume the following three axioms, the last of which is a special case of the axiom of selective convexity introduced by Podinovski (2005).

Axiom 1 (*Feasibility of observed data*) For any $j \in J$, $(X_j, Y_j) \in T$.

Axiom 2 (*Free disposability*) Let $(X, Y) \in T$. Consider any $(\tilde{X}, \tilde{Y}) = (\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ whose subvectors \tilde{X}^R and \tilde{Y}^R satisfy inequalities (1). Let $\tilde{X} \ge X$ and $\tilde{Y} \le Y$. Then $(\tilde{X}, \tilde{Y}) \in T$.

Axiom 3 (Selective convexity) Let $(\tilde{X}, \tilde{Y}) \in T$ and $(\hat{X}, \hat{Y}) \in T$. Assume that $\tilde{X}^R = \hat{X}^R$ and $\tilde{Y}^R = \hat{Y}^R$. Then $\gamma(\tilde{X}, \tilde{Y}) + (1 - \gamma)(\hat{X}, \hat{Y}) \in T$, for any $\gamma \in [0, 1]$.

DMU	Volume input 1	Ratio input 2	Volume output 1	Ratio output 2 (%)
A	6	0.5	4	90
В	3	0.3	1	70
С	4	0.5	2	70

Table 1 An illustration to the notion of R-convex combination

Axiom 3 reflects the fact that, although we cannot generally assume that convex combinations of DMUs in the presence of ratio data remain in technology T, we can nevertheless do so provided the combined DMUs have identical ratio inputs and outputs.²

Importantly, a combination of Axioms 2 and 3 also allows us to form convex combinations of DMUs that have *different* ratio inputs and outputs. This is illustrated by the following example.

Example 1 Consider DMUs A and B shown in Table 1, which are in some technology T with two inputs and two outputs. Input 1 and Output 1 are volume measures, Input 2 is a ratio measure (proportion), and Output 2 is a ratio measure (percentage). We assume that technology T satisfies Axioms 1-3.

Suppose we wish to form a convex combination of DMUs *A* and *B* taken with the weights 1/3 and 2/3, respectively. Note that we cannot use Axiom 3 directly, because the ratio input and output of DMUs *A* and *B* are not identical. Instead, we first employ Axiom 2 and reduce the ratio output of DMU *A* from 90 to 70%, to match the ratio output of DMU *B*. Similarly, we use Axiom 2 to raise the ratio input of DMU *B* from 0.3 to 0.5, to make it equal to the ratio input of DMU *A*.

The resulting DMUs can be stated as $A^* = (6, 0.5, 4, 70\%)$ and $B^* = (3, 0.5, 1, 70\%)$. By Axiom 2, both DMUs A^* and B^* are in technology *T*. Because the ratio input and output of DMUs A^* and B^* are identical, by Axiom 3, any convex combination of these DMUs is in technology *T*. In particular, using the weights 1/3 and 2/3 for DMUs A^* and B^* , respectively, we obtain DMU *C* shown in Table 1.

The above example shows that, although we cannot form convex combinations of DMUs with different ratio inputs and outputs, we can still form their ratio-convex (R-convex) combinations as defined by Olesen et al. (2017). Namely, in a R-convex combination of DMUs, the volume inputs and outputs form conventional convex combinations, while the ratio inputs are taken at their maximum levels across all combined DMUs, and the ratio outputs are taken at their minimum levels. In other words, the weighted average used for volume measures is replaced by the operations of maximum for the ratio inputs and minimum for the ratio outputs.

In line with *the minimum extrapolation principle* used by Banker et al. (1984) and Olesen et al. (2015) give the following definition:

Definition 1 The R-VRS technology $T_{\text{VRS}}^{\text{R}}$ is the intersection of all technologies (sets) $T \subset \mathbb{R}^{m}_{+} \times \mathbb{R}^{s}_{+}$ that satisfy Axioms 1–3.

Following Olesen et al. (2015), technology T_{VRS}^{R} can equivalently be stated as follows.

² For example, assume that two schools of different sizes have the same proportion p of good passes in exams, which is a ratio output. Then, regardless of the underlying (yet unknown to the analyst) numerators and denominators of such proportions, any convex combination of the two schools will always have the same proportion p of good passes.

Table 2DMUs A and B inExample 2	DMU	Volume input	Volume output	Ratio output
•	Α	1	1	0.25
	В	2	3	1.5

Theorem 1 Technology $T_{\text{VRS}}^{\text{R}}$ is the set of all DMUs $(X, Y) = (X^V, X^R, Y^V, Y^R) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ for which there exists a vector $\lambda \in \mathbb{R}^n$ such that

$$\sum_{j \in I} \lambda_j Y_j^V \ge Y^V, \tag{2a}$$

$$\sum_{j \in J} \lambda_j X_j^V \le X^V, \tag{2b}$$

$$\lambda_j \left(Y_j^R - Y^R \right) \ge \mathbf{0}, \quad \forall j \in J,$$
(2c)

$$\lambda_j \left(X_j^R - X^R \right) \le \mathbf{0}, \quad \forall j \in J,$$
 (2d)

$$\mathbf{1}^{\top} \lambda = 1, \tag{2e}$$

$$X^{R} \le X^{R}, \tag{2f}$$

$$Y^R \le \bar{Y}^R, \tag{2g}$$

$$\lambda \ge \mathbf{0}. \tag{2h}$$

In this statement of technology $T_{\text{VRS}}^{\text{R}}$, the first two vector inequalities (2a) and (2b) describe conventional convex combinations of volume inputs and outputs of the observed DMUs taken with the weights $\lambda_i \ge 0$, $j \in J$, that add up to 1 as in equality (2e).

To see the role of inequalities (2c) and (2d), restate them as follows:

if
$$\lambda_j > 0$$
 then $Y_j^R \ge Y^R$ and $X_j^R \le X^R$, $\forall j \in J$. (3)

Conditions (3) imply that an observed DMU $(X_j, Y_j), j \in J$, may be used in the convex combination of volume inputs and outputs in constraints (2a) and (2b) with a positive λ_j only if DMU (X_j, Y_j) is not worse than the DMU (X, Y) on all ratio inputs and outputs.

It is clear that technology $T_{\text{VRS}}^{\text{R}}$ stated by conditions (2) consists of all R-convex combinations of observed DMUs and all DMUs outperformed (dominated) by them, subject to the upper bounds (1) on the ratio measures. This is similar to the conventional VRS technology which includes all convex combinations of observed DMUs and all DMUs outperformed by them.

It is worth noting that $T_{\text{VRS}}^{\text{R}}$ is generally not a convex set, but is a closed set (Olesen et al. 2015).

Example 2 Consider DMUs A and B with one volume input, one volume output and one ratio output as shown in Table 2. Fig. 1 shows technology $T_{\text{VRS}}^{\text{R}}$ generated by these two DMUs. Observe that this technology does not include convex combinations of DMUs A and B because they have different levels of the ratio output. (The line AB is not included in the technology.) Instead, the technology includes the line AD which represents all R-convex combinations of DMUs A and B. On this line, the ratio output is taken at the minimum level 0.25 of the ratio outputs of DMUs A and B, while the volume inputs and outputs of DMUs A and B form standard convex combinations.



Fig. 1 The R-VRS technology generated by DMUs A and B in Example 2

Also note that the technology in Fig. 1 satisfies Axioms 1–3. In particular, Axiom 3 of selective convexity is satisfied because, for each level of ratio output, the corresponding section of technology $T_{\text{VRS}}^{\text{R}}$ is a convex set. For example, the section corresponding to the output level 0.25 of DMU *A* is the convex polyhedron *KCADE*. The section corresponding to the output level 1.5 of DMU *B* is the polyhedron *HFBG*. However, the whole technology $T_{\text{VRS}}^{\text{R}}$ is not a convex set.

Remark 1 Technology $T_{\text{VRS}}^{\text{R}}$ may be seen as a generalization of several technologies. Suppose that there are no ratio inputs and outputs. In this case the inequalities (2c), (2d), (2f) and (2g) are omitted and $T_{\text{VRS}}^{\text{R}}$ is the conventional VRS technology of Banker et al. (1984). If there are no volume inputs and outputs then $T_{\text{VRS}}^{\text{R}}$ is free disposal hull (FDH) of Deprins et al. (1984), with the additional upper bounds (1) on the ratio measures. If we have only volume and ratio inputs and volume outputs, i.e., there are no ratio outputs, technology $T_{\text{VRS}}^{\text{R}}$ is similar to model (7) stated by Ruggiero (1996) in which the ratio factor Z characterizes the quality of the environment.

A subtle difference between our approach and the approach of Ruggiero (1996) is that we consider any ratio input and output (including those of environmental nature) as yet another dimension of technology $T_{\text{VRS}}^{\text{R}}$. In contrast, Ruggiero (1996) does not formally include the environmental factor Z as an input of the technology but instead defines the conventional VRS technology $T_{\text{VRS}}(Z)$, for every value of Z treated as a parameter. It is clear that the parametric family of technologies $T_{\text{VRS}}(Z)$ of Ruggiero is the collection of the sections of technology $T_{\text{VRS}}^{\text{R}}$ defined in the volume input and output dimensions for each fixed value of parameter Z.

The treatment of environmental factors (and any other ratio measures) as inputs and outputs of technology $T_{\text{VRS}}^{\text{R}}$ allows us, if required, to account for such factors in the evaluation of efficiency and various scale characteristics. For example, we may explore the question of optimal scale of production and returns to scale with regard to several inputs and outputs, including environmental factors. These possibilities are considered in the general setting in the subsequent sections.

4 Scale efficiency in the R-VRS technology

In this section we follow the approach of Banker (1984) and show how the simultaneous development of the notions of MPSS and scale efficiency in the standard VRS technology could be extended to technology $T_{\text{VRS}}^{\text{R}}$.

4.1 The general setting

To keep the exposition as general as possible, we consider scale properties of the production frontier with respect to some selected *nonempty* subsets of inputs I' and outputs O':

$$I' \subseteq I = I^V \cup I^R$$
, $O' \subseteq O = O^V \cup O^R$.

In different applications, the sets I' and O' may represent discretionary inputs and outputs in the sense explored by Banker and Morey (1986) and Golany and Roll (1993). These sets may also include measures that are not discretionary, such as certain exogenous factors.³

For example, consider a scenario in which the policy maker uses a DEA model with several volume and ratio inputs and outputs for the assessment of school efficiency. Suppose that this model uses the percentage of families from the higher socio-economic background as an exogenously fixed *non-discretionary* ratio input x' and the percentage of school graduates going to university y' as a *discretionary* ratio output representing quality of education. Suppose that the policy maker is interested in the relationship between these two factors on the production frontier (i.e., among the efficient schools only) while keeping all the other inputs and outputs constant. This leads to the question of optimal scale and returns to scale defined in the selective (partial) sense, for which $I' = \{x'\}$ and $O' = \{y'\}$, even though the input x' is not discretionary.

For a DMU $(X_o, Y_o) \in T_{\text{VRS}}^{\text{R}}$, define vectors $X_o(\varphi)$ and $Y_o(\psi)$, where $\varphi, \psi \ge 0$ are scaling factors, as follows:

$$X_{oi}(\varphi) = \begin{cases} \varphi X_{oi}, & \text{if } i \in \mathbf{I}', \\ X_{oi}, & \text{if } i \in \mathbf{I} \backslash \mathbf{I}', \end{cases} \quad Y_{or}(\psi) = \begin{cases} \psi Y_{or}, & \text{if } r \in \mathbf{O}', \\ Y_{or}, & \text{if } r \in \mathbf{O} \backslash \mathbf{O}'. \end{cases}$$
(4)

The output radial efficiency (technical efficiency) of DMU (X_o, Y_o) measured with respect to the selected subset of outputs O' is defined as the inverse of the optimal value of the following program (in its statement, for consistency of notation maintained throughout this paper, we change variable ψ to η):

$$\tilde{\eta} = \max \eta$$
subject to $(X_o, Y_o(\eta)) \in T_{\text{VRS}}^{\text{R}}, \ \eta \ge 0.$
(5)

Assessing the efficiency of DMU (X_o , Y_o) by program (5) is straightforward. This requires replacing the DMU (X, Y) in conditions (2) by DMU (X_o , $Y_o(\eta)$), and maximizing η subject to the resulting conditions (Olesen et al. 2015, 2017).

³ Olesen et al. (2015) consider a similar setting for the assessment of input and output radial efficiency of DMUs with respect to the selected subsets of inputs and outputs. The authors outline several scenarios in which such a selective approach may be required.

4.2 The partial cone extension of the R-VRS technology

In this section we consider the partial cone extension of technology $T_{\text{VRS}}^{\text{R}}$ defined with respect to the selected input and output sets I' and O'. This cone extension is useful in exploring the notions of MPSS and scale efficiency in $T_{\text{VRS}}^{\text{R}}$ undertaken in subsequent sections.

Define the *partial* cone extension C(I', O') of technology T_{VRS}^{R} as follows:

$$C(\mathbf{I}',\mathbf{O}') = \left\{ (X,Y) \in \mathbb{R}^{m+s} \mid \exists (\tilde{X},\tilde{Y}) \in T_{\text{VRS}}^{\text{R}}, \alpha \ge 0 : (X,Y) = (\tilde{X}(\alpha),\tilde{Y}(\alpha)) \right\}, \quad (6)$$

where the DMU $(\tilde{X}(\alpha), \tilde{Y}(\alpha))$ is as defined by (4) with $\varphi = \psi = \alpha$. The cone C(I', O') includes all partially scaled DMUs $(X(\alpha), Y(\alpha))$ obtained from all DMUs $(X, Y) \in T_{\text{VRS}}^{\text{R}}$, for all $\alpha \ge 0$. If I' = I and O' = O, the cone C(I', O') = C(I, O) is the full cone extension of technology $T_{\text{VRS}}^{\text{R}}$.

It is straightforward to show that the cone C(I', O') is generally not a closed set.⁴ Define its closure as

$$\bar{C}(\mathbf{I}',\mathbf{O}') = \operatorname{cl} C(\mathbf{I}',\mathbf{O}').$$
(7)

The next result provides an explicit statement of the closed cone $\bar{C}(I', O')$. It is proved under the following mild assumption about all observed DMUs.

Assumption 1 For each $j \in J$, there exists an $i \in I'$ (generally different for different j) such that $X_{ji} > 0$.

Theorem 2 Let Assumption 1 be true. Then the closed partial cone extension $\overline{C}(I', O')$ of technology $T_{\text{VRS}}^{\text{R}}$ is the set of all DMUs $(X, Y) = (X^V, X^R, Y^V, Y^R) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ for which there exist a vector $\lambda \in \mathbb{R}^n$ and a scalar σ such that

$$\sum_{i \in J} \lambda_j \sigma Y_{jr}^V \ge Y_r^V, \quad \forall r \in \mathcal{O}^V \cap \mathcal{O}',$$
(8a)

$$\sum_{j \in I} \lambda_j Y_{jr}^V \ge Y_r^V, \quad \forall r \in \mathcal{O}^V \backslash \mathcal{O}',$$
(8b)

$$\sum_{i \in I} \lambda_j \sigma X_{ji}^V \le X_i^V, \quad \forall i \in \mathbf{I}^V \cap \mathbf{I}',$$
(8c)

$$\sum_{i \in I} \lambda_j X_{ji}^V \le X_i^V, \quad \forall i \in \mathbf{I}^V \backslash \mathbf{I}', \tag{8d}$$

$$\lambda_j(\sigma Y_{jr}^R - Y_r^R) \ge 0, \quad \forall j \in J, \quad \forall r \in \mathcal{O}^R \cap \mathcal{O}',$$
(8e)

$$\lambda_j (Y_{jr}^R - Y_r^R) \ge 0, \quad \forall j \in J, \quad \forall r \in \mathcal{O}^R \backslash \mathcal{O}',$$
(8f)

$$\lambda_j(\sigma X_{ji}^R - X_i^R) \le 0, \quad \forall j \in J, \quad \forall i \in \mathbf{I}^R \cap \mathbf{I}',$$
(8g)

$$\lambda_j (X_{ji}^R - X_i^R) \le 0, \quad \forall j \in J, \quad \forall i \in \mathbf{I}^R \setminus \mathbf{I}',$$
(8h)

$$\mathbf{1}^{\top} \boldsymbol{\lambda} = 1, \tag{8i}$$

⁴ For example, consider the VRS technology T'_{VRS} generated by the single DMU A = (1, 1, 5) whose first two components are volume inputs and the third component is a volume output. This technology is special case of technology T^{R}_{VRS} . Let I' = I and O' = O. Consider DMU B = (1, 0, 0). Because technology T'_{VRS} does not include points on the ray { $(\alpha, 0, 0) | \alpha \ge 0$ }, DMU B is not in the (full) cone C(I, O). However, B is is in the CRS technology T'_{CRS} generated by DMU A which is the closure of C(I, O). (Note that B is dominated by the origin $(0, 0, 0) \in T'_{CRS}$ and is included in T'_{CRS} by the axiom of free disposability). Therefore, the cone C(I, O) is not a closed set.

$$Y_r^R \le \sigma \bar{Y}_r^R, \quad \forall r \in \mathcal{O}^R \cap \mathcal{O}', \tag{8j}$$

$$Y_r^R < \bar{Y}_r^R, \quad \forall r \in \mathcal{O}^R \backslash \mathcal{O}', \tag{8k}$$

$$X_i^R < \sigma \bar{X}_i^R, \quad \forall i \in \mathbf{I}^R \cap \mathbf{I}', \tag{81}$$

$$X_i^R \le \bar{X}_i^R, \quad \forall i \in \mathbf{I}^R \backslash \mathbf{I}', \tag{8m}$$

$$\lambda \ge \mathbf{0}, \sigma \ge 0. \tag{8n}$$

It is clear that conditions (8) are closely related to conditions (2). The difference is that, in (8), the scaling factor σ is attached to all volume and ratio inputs $i \in I'$ and outputs $r \in O'$ with respect to which we define the cone extension, and the remaining inputs and outputs in the sets I \ I' and O \ O' are kept fixed. As a result, the closed *partial* cone extension $\overline{C}(I', O')$ of technology $T_{\text{VRS}}^{\text{R}}$ includes all partially scaled R-convex combinations of the observed DMUs, where only the inputs and outputs in the sets I' and O' are scaled by $\sigma \ge 0$, and the remaining inputs $i \in I \setminus I'$ and $r \in O \setminus O'$ are not scaled.

4.3 The most productive scale size in the R-VRS technology

Recall that Banker (1984) introduces the notion of MPSS evaluated with respect to the entire vectors of inputs and outputs in the VRS technology T_{VRS} (which, according to Remark 1, can be viewed as a special case of the R-VRS technology $T_{\text{VRS}}^{\text{R}}$). Namely, let DMU $(X_o, Y_o) \in T_{\text{VRS}}$, and let $\varphi > 0$ and $\psi > 0$ be the scaling factors for the input and output vectors, respectively. Consider the parametric set of all DMUs $(\varphi X_o, \psi Y_o) \in T_{\text{VRS}}$. All such DMUs have the same structure of the input and output vectors as the original DMU (X_o, Y_o) . Then DMU (X_o, Y_o) is at MPSS if, for any φ and ψ such that $(\varphi X_o, \psi Y_o) \in T_{\text{VRS}}$, we have $\psi/\varphi \leq 1$. In other words, DMU (X_o, Y_o) is at MPSS if it maximizes the average productivity ψ/φ among all DMUs $(\varphi X_o, \psi Y_o) \in T_{\text{VRS}}$.

Below we follow the same approach and introduce the notion of MPSS in technology T_{VRS}^{R} evaluated with respect to the subsets I' and O'. Consider the program

$$\theta^* = \sup (\psi/\varphi)$$

subject to $(X_o(\varphi), Y_o(\psi)) \in T_{\text{VRS}}^{\text{R}}, \ \varphi > 0, \ \psi \ge 0,$ (9)

where the vectors $X_o(\varphi)$ and $Y_o(\psi)$ are defined as in (4).

If the sets I' and O' include all inputs and outputs (i.e., if I' = I and O' = O), the condition $(X_o(\varphi), Y_o(\psi)) \in T_{\text{VRS}}^{\text{R}}$ of program (9) is restated as $(\varphi X_o, \psi Y_o) \in T_{\text{VRS}}^{\text{R}}$. In this case, similar to the standard definition of Banker (1984), program (9) maximizes the average productivity ψ/φ among all DMUs stated in the form $(\varphi X_o, \psi Y_o) \in T_{\text{VRS}}^{\text{R}}$, i.e., preserving the input and output structures of the DMU (X_o, Y_o) under the consideration.

In the general case, program (9) maximizes the ratio ψ/φ which is interpretable as the ratio of the quantity ψ of the subvector of selected outputs Y_{or} , $r \in O'$, to the quantity φ of the subvector of selected inputs X_{oi} , $i \in I'$, while keeping the remaining inputs and outputs constant.

Because $\varphi = \psi = 1$ is a feasible solution of program (9), we always have $\theta^* \ge 1$.

Definition 2 DMU $(X_o, Y_o) \in T_{\text{VRS}}^{\text{R}}$ is at MPSS evaluated with respect to the subsets I' and O' if $\theta^* = 1$.

Let us show that solving program (9) can be replaced by assessment of the *partial* output radial efficiency of DMU (X_o, Y_o) (with respect to the outputs in the set O' only) in the

closed partial cone extension $\bar{C}(I', O')$ of technology $T_{\text{VRS}}^{\text{R}}$ whose statement is obtained by Theorem 2. This result is formally stated and proved as Theorem 3 below. Let us first provide its intuitive explanation.

Consider program (9). Note that we can substitute technology $T_{\text{VRS}}^{\text{R}}$ in its constraints by its cone extension C(I', O'). Indeed, for any feasible solution $\langle \varphi, \psi \rangle$ of program (9), changing $T_{\text{VRS}}^{\text{R}}$ to C(I', O') adds the full ray $\langle \alpha \varphi, \alpha \psi \rangle, \alpha > 0$, to its feasible region. However, because $\alpha \psi / \alpha \varphi = \psi / \varphi$, this change does not affect the supremum of the objective function.

Furthermore, because the value of the objective function ψ/φ of the resulting program (9) with $T_{\text{VRS}}^{\text{R}}$ replaced by C(I', O') is constant along any ray of feasible solutions $\langle \alpha \varphi, \alpha \psi \rangle$, $\alpha > 0$, we can restrict the feasible set to only one point on each ray, by requiring that $\varphi = 1$. After this normalization of variable φ and renaming variable ψ to η , we obtain the following program:

$$\eta' = \sup \eta$$

subject to $(X_o, Y_o(\eta)) \in C(\mathbf{I}', \mathbf{O}'), \ \eta \ge 0.$ (10)

Taking into account that the cone C(I', O') is generally not a closed set, we replace it in the constraints of program (10) by the closed cone $\overline{C}(I', O')$. This results in a linear program in which we conventionally change the supremum of the objective function to its maximum:

$$\eta^* = \max \eta$$

subject to $(X_o, Y_o(\eta)) \in \overline{C}(I', O'), \ \eta \ge 0.$ (11)

The next result requires an additional assumption about DMU (X_o, Y_o) which should be true in any meaningful application:

Assumption 2 There exists an $r \in O'$ such that $Y_{or} > 0$.

Theorem 3 Let both Assumptions 1 and 2 be true. Then the supremum θ^* of program (9) is equal to the maximum η^* of program (11), and both are attained.

Corollary 1 The DMU $(X_o, Y_o(\eta^*)) \in \overline{C}(I', O')$ at which the maximum in program (11) is attained is also in the cone C(I', O').

It is clear that program (11) which uses the statement (8) of the cone C(I', O') is nonlinear and would be problematic in practical computations. Therefore, Theorem 3 should primarily be of theoretical interest. It shows that the evaluation of MPSS for DMU (X_o , Y_o) in technology $T_{\text{VRS}}^{\text{R}}$ (with respect to the subsets I' and O') is equivalent to the evaluation of its partial output radial efficiency (with respect to the subset O') in the closed partial cone extension $\bar{C}(I', O')$ of technology $T_{\text{VRS}}^{\text{R}}$. In particular, DMU (X_o , Y_o) is at MPSS with respect to I' and O' in $T_{\text{VRS}}^{\text{R}}$ if and only if its partial output radial efficiency in $\bar{C}(I', O')$ evaluated with respect to the subset O' is equal to 1.

Consider the special case in which technology $T_{\text{VRS}}^{\text{R}}$ is the standard VRS technology T_{VRS} of Banker et al. (1984) and the sets I' and O' include all inputs and outputs. Then the closed cone $\bar{C}(I', O')$ is the standard CRS technology T_{CRS} of Charnes et al. (1978). In this case, Theorem 3 becomes a well-known result that DMU_o $\in T_{\text{VRS}}$ is at MPSS if and only if its output radial efficiency η^* in the benchmark CRS technology T_{CRS} is equal to 1.

In Sect. 5, we consider computational approaches to solving program (9). We transform this program to an equivalent form that, depending on the sets I, O, I' and O', either becomes a linear program or can be solved as a mixed integer linear program.

4.4 Overall and scale efficiency in the R-VRS technology

Following the approach of Banker (1984), we interpret the inverse optimal value $1/\theta^*$ of program (9) [or, equivalently, the inverse value $1/\eta^*$ of program (11)] as the overall efficiency $OE(X_o, Y_o)$ of DMU (X_o, Y_o) assessed with respect to the subsets I' and O'.

By Definition 2, DMU (X_o, Y_o) is at MPSS if and only if $OE(X_o, Y_o) = 1$. Otherwise we can decompose the overall efficiency into the product of its technical and scale efficiency components $TE(X_o, Y_o)$ and $SE(X_o, Y_o)$:

$$OE(X_o, Y_o) = TE(X_o, Y_o) \times SE(X_o, Y_o),$$

where $TE(X_o, Y_o) = 1/\tilde{\eta}$ is evaluated by solving program (5). Then $SE(X_o, Y_o) = \tilde{\eta}/\eta^*$ is the scale efficiency of DMU (X_o, Y_o) evaluated with respect to the subsets I' and O'. Because $T_{\text{VRS}}^{\text{R}} \subset \overline{C}(I', O')$, we have $\tilde{\eta} \le \eta^*$ and $SE(X_o, Y_o) \le 1$.

Following Banker (1984), the scale efficiency $SE(X_o, Y_o)$ is interpretable as a measure of divergence from MPSS with respect to the sets I' and O'. Indeed, let for simplicity DMU (X_o, Y_o) be technically efficient but scale inefficient. Then DMU (X_o, Y_o) is not at MPSS (for the selected sets I' and O') and, in program (9), we have $\theta^* > 1$. By Theorem 3, the supremum θ^* is attained at some feasible solution $\langle \varphi^*, \psi^* \rangle$ of program (9), and we have $\theta^* = \psi^*/\varphi^* > 1$.

The DMU ($X_o(\varphi^*)$, $Y_o(\psi^*)$) defined by (4) represents MPSS for DMU (X_o, Y_o). If DMU (X_o, Y_o) scales its inputs in the set I' by the factor φ^* and outputs in the set O' by the factor ψ^* , while keeping the remaining inputs and outputs fixed, its average productivity (measured only with respect to the selected inputs and outputs) will increase by $\psi^*/\varphi^* > 1$. Therefore, the ratio $\psi^*/\varphi^* = 1/SE(X_o, Y_o)$ shows by how much the average productivity of this DMU, measured with respect to the selected inputs and outputs in the sets I' and O', could increase if it were to change these inputs and outputs by the factors φ^* and ψ^* , respectively, to match those of its MPSS. This is similar to the standard notion of MPSS in the VRS technology. Following Banker (1984), the optimal ratio ψ^*/φ^* is interpretable as a measure of divergence of the technically efficient DMU (X_o, Y_o) from its MPSS, if we restrict the scaling of the inputs and outputs to the sets I' and O' only, while keeping the remaining inputs and outputs constant.

This interpretation is illustrated by the following example.

Example 3 Let us refer to technology $T_{\text{VRS}}^{\text{R}}$ in Example 2. Consider the following three scenarios in which the set I' includes the single input but the sets O' are defined differently. Note that both DMUs A and B are technically efficient in all three scenarios.

(i) Let the set O' include both the volume and ratio outputs. Consider assessing the scale efficiency of DMU *A* by program (9). Its unique optimal solution is $\varphi^* = 2$ and $\psi^* = 3$, which corresponds to DMU *P* (see Table 3 and Fig. 2).⁵ DMU *P* uses twice the amount of volume input of DMU *A* but produces three times its vector of outputs. The optimal value of the objective function of program (9) is equal to 3/2, and the scale efficiency of DMU *A* is 2/3. Note that, in the given scenario, DMU *P* represents MPSS for DMU *A*.

⁵ We can consider φ changing from 1 to $+\infty$ ($\varphi = 1$ corresponds to DMU A). For $\varphi \in [1, 2)$, the maximum value of ψ subject to the constraints of program (9) is equal to 1. This corresponds to the line AL in Fig. 2. For $\varphi \ge 2$, the maximum value of ψ changes to 3 and remains constant as φ increases. This corresponds to the ray PQ. The maximum ratio ψ/φ is attained at $\varphi = 2$ and $\psi = 3$, i.e., at point P.

Table 3DMUs representingMPSS in Example 3	DMU	Volume input	Volume output	Ratio output
	D	2	3	0.25
	Р	2	3	0.75
	V	2	1	1.5



Fig. 2 MPSS in the R-VRS technology evaluated with respect to different sets of outputs O'

Similarly, in the case of DMU *B*, the unique optimal solution to program (9) is $\varphi^* = \psi^* = 1$. Therefore, DMU *B* is at MPSS.⁶

- (ii) Let the set O' include only the volume output but not the ratio output. In this case, program (9) assesses the scale efficiency of DMU *A* by keeping its ratio output constant, i.e., by restricting the evaluation to the section of technology *KCADE*. In this case the optimal solution is $\varphi^* = 2$ and $\psi^* = 3$. This means that the scale efficiency of DMU *A* is equal to 2/3 and its MPSS is DMU *D*. Similarly, the scale efficiency of DMU *B* is assessed on the section *HFBG*, and DMU *B* is at MPSS.
- (iii) Let the set O' include only the ratio output but not the volume output. To evaluate the scale efficiency of DMU A, we search among all DMUs on the broken line ALVW. The highest ratio of the ratio output to the volume input is achieved at DMU V, which corresponds to $\varphi^* = 2$ and $\psi^* = 6$ in program (9). Therefore, in this scenario, the scale efficiency of DMU A is $\varphi^*/\psi^* = 1/3$, and DMU V represents MPSS for DMU A. A similar investigation shows that DMU B is at MPSS.

4.5 The cone extension and the CRS extension of the R-VRS technology

The cone and CRS extensions of the standard VRS technology are the same sets, and both can be used as the reference technologies in the evaluation of MPSS and scale efficiency of the DMUs. The purpose of this section is to demonstrate that the same identity does not

⁶ For any $\varphi \ge 1$ in program (9), the maximum value of ψ is equal to 1 which corresponds to the ray *BG*. For $\varphi \in [0.5, 1)$, the maximum value of ψ is equal to 1/6, which corresponds to the line *UZ*, where U = (1, 0.5, 0.25). The values of $\varphi < 0.5$ are infeasible.

hold for the R-VRS technology. Namely, its CRS extension is generally different from its cone extension, and it is the latter that we employ in the evaluation of MPSS and other scale characteristics.

In order to demonstrate this difference and avoid excessive technicalities, for the discussion in this section, we assume that I' = I and O' = O. We also assume that the bounds (1) are not specified.

As noted in Remark 1, if all inputs and outputs are volume measures, technology $T_{\text{VRS}}^{\text{R}}$ is the standard VRS technology T_{VRS} of Banker et al. (1984). In this case, program (9) represents the standard approach of Banker (1984) for the assessment of MPSS in the VRS technology. Further, the closed cone extension of technology T_{VRS} coincides with its CRS extension T_{CRS} , which is the CRS technology of Charnes et al. (1978). This implies that the evaluation of MPSS for DMU (X_o , Y_o) by solving program (9) (in the case in which all inputs and outputs are volume measures) is equivalent to the assessment of output radial efficiency of this DMU in the CRS technology T_{CRS} .

Below we show that the same equivalence does not hold in the case of technology $T_{\text{VRS}}^{\text{R}}$. To demonstrate this, we use the CRS extension of technology $T_{\text{VRS}}^{\text{R}}$ introduced by Olesen et al. (2015) in which the ratio inputs and outputs are of the proportional type with no bounds. Such inputs and outputs are assumed to be scalable in the same proportion as the volume measures. Following Olesen et al. (2015), we denote such technology $T_{\text{VRS}}^{\text{P}}$. This technology is defined (in the sense of the minimum extrapolation principle) by Axioms 1–3 and the following additional axiom:

Axiom 4 (*Proportionality*) Let $(X^V, X^R, Y^V, Y^R) \in T$. Then, for all scaling factors $\alpha \ge 0$, $(\alpha X^V, \alpha X^R, \alpha Y^V, \alpha Y^R) \in T$.

By Theorem 2 in Olesen et al. (2015), technology T_{CRS}^P is the set of all DMUs $(X, Y) = (X^V, X^R, Y^V, Y^R) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ for which there exist vectors $\lambda, \sigma \in \mathbb{R}^n$ such that

$$\sum_{j \in J} \lambda_j \sigma_j Y_j^V \ge Y^V,$$

$$\sum_{j \in J} \lambda_j \sigma_j X_j^V \le X^V,$$

$$\lambda_j (\sigma_j Y_j^R - Y^R) \ge \mathbf{0}, \quad \forall j \in J,$$

$$\lambda_j (\sigma_j X_j^R - X^R) \le \mathbf{0}, \quad \forall j \in J,$$

$$\mathbf{1}^\top \lambda = 1,$$

$$\lambda, \sigma \ge \mathbf{0}.$$
(12)

The full closed cone extension $\overline{C}(I, O)$ of technology $T_{\text{VRS}}^{\text{R}}$ and its R-CRS extension $T_{\text{CRS}}^{\text{P}}$ are in general different sets. Indeed, the former is stated by conditions (8) in which I' = I and O' = O. In these conditions, all observed DMUs $(X_j, Y_j), j \in J$, are scaled by the same single factor σ . In contrast, in (12), all observed DMUs are scaled independently by generally different factors $\sigma_j, j \in J$. Therefore, we have $\overline{C}(I, O) \subseteq T_{\text{CRS}}^{\text{P}}$. Example 4 considered below and a further example in Appendix B show that, generally, this embedding is not an equality.

To facilitate the graphical illustration in the next example, we first establish the following useful result which is true if the specification of technology T_{CRS}^{P} includes only a single ratio measure. This measure can be a ratio input or a ratio output, but not both. This result is not valid if the technology incorporates more than one ratio measure.

Table 4

DMUs in Example 4	DMU	x ^V	<i>x</i> ^{<i>R</i>}	y ^V
	A	3	1	2
	В	1	2	1.5
	С	3	2	2
	D = 0.5(B + C)	2	2	1.75



Fig. 3 Technology $T_{\text{VRS}}^{\text{R}}$ in Example 4

Theorem 4 Let technology T_{CRS}^{P} include a single ratio measure, i.e., let the union of the sets $I^{R} \cup O^{R}$ be a singleton. Also, let this ratio measure be strictly positive for every observed DMU. (E.g., in the case of single ratio input X^{R} , we assume that $X_{j}^{R} > 0$, for all j = 1, ..., n.) Then technology T_{CRS}^{P} is the standard CRS technology T_{CRS} generated by the same set of observed DMUs.

As established in Sect. 4.3, the maximum average productivity that DMU (X_o, Y_o) achieves at its MPSS in technology $T_{\text{VRS}}^{\text{R}}$, as represented by the maximum of the objective function ψ/φ of program (9), is equal to the average productivity of its output radial projection in the closed cone extension $\bar{C}(I, O)$. The following example illustrates the case in which the maximum average productivity for DMU (X_o, Y_o) with $T_{\text{VRS}}^{\text{R}}$ (or $\bar{C}(I, O)$) as the benchmark technology falls below the average productivity in technology $T_{\text{CRS}}^{\text{P}}$.

Example 4 The example involves four DMUs A, B, C and D with one volume input x^V , one proportional ratio input x^R and one volume output y^V , as shown in Table 4. Technology T_{VRS}^R generated by DMUs A, B and C is shown in Fig. 3. It is not convex

Technology $T_{\text{VRS}}^{\text{K}}$ generated by DMUs *A*, *B* and *C* is shown in Fig. 3. It is not convex and consists of the two shaded parts. Note that the ratio input x^R of the DMUs *B* and *C* is equal. By Axiom 3 of selective convexity, the line segment *BC* is included in technology $T_{\text{VRS}}^{\text{R}}$. Because the ratio input x^R of the DMUs *A* and *B* is different, the line segment *AB* is not included in this technology.

Let us evaluate the scale efficiency of DMU *D*. This requires that we consider all DMUs in $T_{\text{VRS}}^{\text{R}}$ stated as $(2\alpha, 2\alpha, y^{V})$ and identify the largest ratio y^{V}/α among all such DMUs.

Consider the two-dimensional piecewise linear section of technology $T_{\text{VRS}}^{\text{R}}$ with the input mix described parametrically as $(2\alpha, 2\alpha)$. In other words, we measure input by $\alpha \ge 0$. The



Fig. 4 The (α, y^V) -diagram of the two-dimensional sections of technologies $T_{\text{VRS}}^{\text{R}}$, $T_{\text{CRS}}^{\text{P}}$ and the closed cone $\bar{C}(I, O)$

mix $(\alpha, y^V) = (1, 1.75)$ is feasible and corresponds to DMU *D*. This is illustrated in Fig. 4 in which the red piecewise linear line corresponds to the line $\overline{D}DLM$ in Fig. 3. The closed cone $\overline{C}(I, O) = \overline{C}(\{1, 2\}, \{1\})$ generated by technology $T_{\text{VRS}}^{\text{R}}$ according to (6) is represented by the ray OK.

It is clear that DMU *D* is scale efficient. It achieves the maximum average productivity equal to $y/\alpha = 1.75/1$ among all DMUs stated in the form $(2\alpha, 2\alpha, y^V)$ and plotted on the red line in Fig. 4. Expressed differently, DMU *D* is also located on the ray *OK*. This ray represents the upper boundary of the closed cone extension $\bar{C}(I, O)$ of $T_{\text{VRS}}^{\text{R}}$. The R-CRS technology $T_{\text{CRS}}^{\text{P}}$ is shown as the shaded area in Fig. 5. In this example, by

The R-CRS technology T_{CRS}^{P} is shown as the shaded area in Fig. 5. In this example, by Theorem 4, it coincides with the standard CRS technology T_{CRS} generated by the observed DMUs A, B and C. The cone spanned by DMUs A and B of dimension 2 is a facet in T_{CRS}^{P} .⁷ DMU D is located below this facet and is an interior point of technology T_{CRS}^{P} .

DMU *D* is located below this facet and is an interior point of technology T_{CRS}^{P} . Consider the two-dimensional piecewise linear section of T_{CRS}^{P} with the input mix fixed to (2, 2) of DMU *D*. The possible increase in the observed average product for *D* (equal to 1.75) to be obtained with T_{CRS}^{P} as benchmark can now be estimated by the maximal expansion of the current output (equal to 1.75) with inputs fixed at the level (2, 2).

Straightforward calculations show that the output projection of DMU D on the boundary of technology T_{CRS}^{P} is DMU D' with the input mix equal to (2, 2) and $y^{V} = 2$. Using the

⁷ It is also straightforward to show that the segment *AB* is in T_{CRS}^{P} independently of Theorem 4. Both DMUs $\left(x_{A}^{V}, x_{A}^{R}, y_{A}^{V}\right) = (3, 1, 2)$ and $\left(x_{B}^{V}, x_{B}^{R}, y_{B}^{V}\right) = (1, 2, 1.5)$ belong to T_{CRS}^{P} . Applying Axiom 4 of proportionality to DMU *B* with $\alpha = 0.5$, we obtain DMU $B^{*} = (0.5, 1, 0.75) \in T_{CRS}^{P}$. By Axiom 3 of selective convexity, any convex combination of DMUs *A* and B^{*} (which have the same ratio input $x^{R} = 1$), belongs to T_{CRS}^{P} . The facet spanned by these convex combinations is identical to the facet spanned by the line segment *AB*.



Fig. 5 Technology T_{CRS}^{P} in Example 4

result of Theorem 4, we identify DMU D' as the radial projection of DMU D in the standard output-oriented CRS model with DMUs A, B and C as the three observed DMUs.

The maximal output expansion is shown by the movement of D to D' in Fig. 5. DMU D' is located on the facet spanned by A and B. Point D is MPSS and SE(D) = 1 with $T_{\text{VRS}}^{\text{R}}$ as the benchmark technology, while D is not at MPSS with $T_{\text{CRS}}^{\text{P}}$ as the benchmark technology. The output radial improvement factor for DMU D is now $\theta = 2/1.75 \approx 1.143$ with $T_{\text{CRS}}^{\text{P}}$ as benchmark, because the maximal average product for DMU D increases from 1.75 with $T_{\text{VRS}}^{\text{R}}$ as benchmark to 2 with $T_{\text{CRS}}^{\text{P}}$ as benchmark.

5 Computation of scale efficiency

As highlighted in Sect. 4, the evaluation of MPSS and overall and scale efficiency by nonlinear programs (9) or (11) is generally problematic. In this section we obtain an equivalent statement of these programs whose use in practical computations is straightforward.

Below, when referring to program (9), we assume that its constraint $(X_o(\varphi), Y_o(\psi)) \in T_{\text{VRS}}^{\text{R}}$ is replaced by conditions (2) in which the DMU (X, Y) is changed to $(X_o(\varphi), Y_o(\psi))$. This program is stated in terms of the variable vector λ and scalars φ and ψ . Similarly, we restate program (11) using the statement of the closed partial cone $\overline{C}(I', O')$ by Theorem 2. (The full statement of the resulting program (11) is shown as program (28) in the proof of Theorem 5.) We now use the substitution $\lambda' = \lambda \sigma$. Suppressing the prime symbol, we state the resulting program as

$$\hat{\eta} = \max \ \eta \tag{13a}$$

subject to

$$\sum_{i \in I} \lambda_j Y_{jr}^V \ge \eta Y_{or}^V, \quad \forall r \in \mathcal{O}^V \cap \mathcal{O}',$$
(13b)

$$\sum_{j \in J} \lambda_j Y_{jr}^V \ge \sigma Y_{or}^V, \quad \forall r \in \mathbf{O}^V \backslash \mathbf{O}',$$
(13c)

Deringer

$$\sum_{i \in J} \lambda_j X_{ji}^V \le X_{oi}^V, \quad \forall i \in \mathbf{I}^V \cap \mathbf{I}',$$
(13d)

$$\sum_{i \in I} \lambda_j X_{ji}^V \le \sigma X_{oi}^V, \quad \forall i \in \mathbf{I}^V \backslash \mathbf{I}',$$
(13e)

$$\lambda_j(\sigma Y_{jr}^R - \eta Y_{or}^R) \ge 0, \quad \forall j \in J, \quad \forall r \in \mathcal{O}^R \cap \mathcal{O}', \tag{13f}$$

$$\lambda_j (Y_{jr}^{\kappa} - Y_{or}^{\kappa}) \ge 0, \quad \forall j \in J, \quad \forall r \in \mathcal{O}^{\kappa} \setminus \mathcal{O}',$$
(13g)

$$\lambda_j(\sigma X_{ji}^R - X_{oi}^R) \le 0, \quad \forall j \in J, \quad \forall i \in \mathbf{I}^R \cap \mathbf{I}',$$
(13h)

$$\lambda_j (X_{ji}^R - X_{oi}^R) \le 0, \quad \forall j \in J, \quad \forall i \in \mathbf{I}^R \setminus \mathbf{I}',$$
(13i)

$$\mathbf{1}^{\top}\boldsymbol{\lambda} = \boldsymbol{\sigma},\tag{13j}$$

$$\eta Y_{or}^R \le \sigma \bar{Y}_r^R \qquad \forall r \in \mathcal{O}^R \cap \mathcal{O}', \tag{13k}$$

$$Y_{or}^{R} \le \bar{Y}_{r}^{R}, \quad \forall r \in \mathcal{O}^{R} \backslash \mathcal{O}', \tag{131}$$

$$X_{oi}^{R} \le \sigma \bar{X}_{i}^{R}, \quad \forall i \in \mathbf{I}^{R} \cap \mathbf{I}', \tag{13m}$$

$$X_{oi}^{R} \leq \bar{X}_{i}^{R}, \quad \forall i \in \mathbf{I}^{R} \backslash \mathbf{I}', \tag{13n}$$

$$\lambda \ge \mathbf{0}, \ \sigma, \eta \ge 0. \tag{130}$$

In program (13), the bounds (131) and (13n) do not restrict decision variables. They are included only to show the relationship of program (13) with the statement (8) of the closed cone $\bar{C}(I', O')$ and can be omitted as redundant in actual computations.

The next result establishes a one-to-one correspondence between the optimal solutions of programs (9) and (13). We prove it under the following stronger variant of Assumption 2:

Assumption 3 Either there exists an $r' \in O^V \cap O'$ such that $Y_{or'} > 0$, or there exists an $r'' \in O^R \cap O'$ such that both $Y_{or''} > 0$ and the upper bound $\overline{Y}_{r''}^R$ in the corresponding constraint (13k) is finite. (Such constraint needs to be specified and should not be omitted.⁸)

Clearly, a simple sufficient, but not necessary, condition that guarantees that Assumption 3 is satisfied, is that all outputs of DMU (X_o , Y_o) are strictly positive and, additionally, all ratio outputs have a specified finite upper bound (1).

Theorem 5 Let Assumptions 1 and 3 be true. Then the following statements are true:

- (i) The maximum value η̂ of program (13) is attained and equal to the supremum θ* of program (9).
- (ii) In any optimal solution $\langle \lambda, \sigma, \eta \rangle$ to program (13), we have $\sigma > 0$.
- (iii) Solution (λ̂, η̂, σ̂) is optimal in program (13) if and only if the solution (λ', φ', ψ'), where λ' = λ̂/σ̂, φ' = 1/σ̂ and ψ' = η̂/σ̂, is optimal in program (9).

To see the importance of Theorem 5, recall that the evaluation of MPSS and the overall efficiency of DMU (X_o, Y_o) rely on our ability to solve program (9). Any of its optimal solutions $\langle \varphi', \psi' \rangle$ would correspond to the MPSS $(X_o(\varphi'), Y_o(\psi'))$ of DMU (X_o, Y_o) in technology $T_{\text{NRS}}^{\text{R}}$, defined with respect to the selected input and output sets I' and O'. The

⁸ To see the importance of Assumption 3, suppose it is not true. Consider any scenario in which the constraint (13m) is redundant (e.g., the set $I^R \cap I' = \emptyset$ or the bounds for inputs in this set are not specified). Then $\lambda = \mathbf{0}, \sigma = 0$ and any $\eta \ge 0$ is a feasible solution of program (13), and its supremum $\hat{\eta}$ is unbounded.

inverse of the optimal value $1/\theta^*$ is the overall efficiency of the DMU (X_o, Y_o) . Because program (9) is nonlinear, we face obvious computational challenges in its practical use.

Theorem 3 stated in Sect. 4.3 appears to suggest a conventional way to overcome this computational difficulty. Namely, instead of evaluating the maximum average productivity ψ/φ by solving the nonlinear program (9), we may equivalently solve program (11) which evaluates the partial (with respect to the set O') output radial efficiency of DMU (X_o , Y_o) in the closed partial cone extension $\overline{C}(I', O')$ of technology $T_{\text{VRS}}^{\text{R}}$.

Unfortunately, the described conventional approach based on solving program (11) instead of (9) does not resolve all computational problems. First, if we state program (11) in the full extended form (shown in the proof of Theorem 5), we obtain a nonlinear program. Second, even if we solve this program and identify its optimal solution $\langle \lambda^*, \eta^*, \sigma^* \rangle$, Theorem 3 does not tell us how to convert this solution to an optimal solution of program (9) and identify the MPSS of DMU (X_o , Y_o).

The new Theorem 5 resolves the above problems. It shows that the nonlinear program (11) can be equivalently restated as program (13). As discussed in Remark 2 below, the latter program is easy to solve. Furthermore, Theorem 5 establishes a one-to-one correspondence between the optimal solutions of programs (9) and (13), and provides simple formulae that transform an optimal solution of either program to an optimal solution of the other program. This means that, in practice, we can solve only program (13) and subsequently convert its optimal solution to an optimal solution of program (9), thus identifying the MPSS of DMU (X_o, Y_o) .

Remark 2 In practical applications, solving program (13) should be unproblematic. Indeed, if the sets I' and O' do not include ratio inputs and outputs, then the sets $I^R \cap I'$ and $O^R \cap O'$ are empty, the constraints (13f), (13h), (13k) and (13m) are omitted and program (13) becomes a linear program. If at least one ratio input or output is included in the sets I' and O', then program (13) is nonlinear. In this case, the constraints (13f) and (13h) can be restated as "either-or" conditions and further linearized using the "big M" method as discussed in Olesen et al. (2017). This transforms program (13) to a mixed integer linear program.

6 Special cases

Below we consider several special cases of program (13). Recall that, as discussed in Sect. 5, solving this program is equivalent to the identification of MPSS by program (9).

6.1 The model of Banker and Morey (1986)

Suppose that there are no ratio measures, i.e., $I = I^V$ and $O = O^V$. In this case, the inequalities (13f)–(13i) and (13k)–(13n) are omitted from program (13). If the set O' includes all volume outputs but I' does not include all volume inputs, the resulting program (13) is the model of Banker and Morey (1986) in which the inputs in the complementary subset $I^V \setminus I'$ are regarded as non-discretionary. If additionally some (non-discretionary) outputs are permitted and are not included in the set O', program (13) is a generalization of the model of Banker and Morey (1986) provided by Golany and Roll (1993). In both models, the inequalities (13c) and (13e) disallow proportional scaling of the volume measures that are not included in the set I' and O'.

6.2 The partial cone extension of the model of Ruggiero (1996)

Let $I' = I^V$ and $O' = O^V$. This corresponds to an important practical situation in which the ratio inputs and outputs, often representing environmental and quality characteristics, are assumed constant in the evaluation of the technical and scale efficiency (Ruggiero 1996). In the described scenario, conditions (13c), (13e), (13f), (13h) and (13k)–(13n) are omitted, and equality (13j) and variable σ become redundant. Then program (13) is stated as follows:

$$\hat{\eta} = \max \eta$$
subject to
$$\sum_{j \in J} \lambda_j Y_j^V \ge \eta Y_o^V,$$

$$\sum_{j \in J} \lambda_j X_j^V \le X_o^V,$$

$$\lambda_j \left(Y_j^R - Y_o^R\right) \ge \mathbf{0}, \quad \forall j \in J,$$

$$\lambda_j \left(X_j^R - X_o^R\right) \le \mathbf{0}, \quad \forall j \in J,$$

$$\lambda \ge \mathbf{0}, \eta \ge \mathbf{0}.$$
(14)

The technology employed by model (14) allows proportional scaling of the volume inputs and outputs while keeping the ratio measures constant. This corresponds to the special case of the R-CRS technology, denoted T_{CRS}^{F} , in which all ratio inputs and outputs are of the fixed type (Olesen et al. 2015).⁹ Technology T_{CRS}^{F} can be viewed as the partial cone extension (with respect to the volume inputs and outputs only) of technology T_{VRS}^{R} stated by Theorem 1. It can also be seen as the CRS extension of the model of Ruggiero (1996) which allows proportional scaling of the volume measures, while keeping the exogenous ratio measures fixed.

6.3 The full scale efficiency

Let the sets I' and O' include all volume and ratio measures, i.e., let I' = I and O' = O. This case was illustrated by scenario (i) of Example 3. As a practical example, in an application to schools, volume measures may represent the teaching hours, expenditure and the number of pupils. The ratio inputs and outputs may represent the percentage of pupils with good grades on entry and exit, and also the percentage of pupils from the higher socio-economic background. The policy maker may be interested in the full scale characterization of the production frontier, in which case the sets I' and O' include all, volume and ratio, inputs and outputs. Note that, regardless of whether some inputs or outputs are discretionary or non-discretionary, it may still be useful to consider the full scale characterization that takes into account all such measures—see Sect. 4.1.

In the described scenario, program (13) takes on the following simplified form:

$$\eta^* = \max \eta$$
 (15a)
subject to

⁹ As a special case of program (13), program (14) is based on Assumption 3. This means that $Y_o^V \neq \mathbf{0}$. As proved by Olesen et al. (2015), under this assumption, the technology employed by model (14) coincides with technology $T_{\text{CRS}}^{\text{F}}$. Without the assumption that $Y_o^V \neq \mathbf{0}$, the statement of technology $T_{\text{CRS}}^{\text{F}}$ is more complex.

$$\sum_{i \in I} \lambda_j X_j^V \le X_o^V, \tag{15c}$$

$$\lambda_j(\sigma Y_j^R - \eta Y_o^R) \ge \mathbf{0}, \qquad \forall j \in J,$$
(15d)

$$\lambda_j(\sigma X_j^R - X_o^R) \le \mathbf{0}, \quad \forall j \in J,$$
(15e)

$$\mathbf{1}^{\mathsf{T}}\boldsymbol{\lambda} = \boldsymbol{\sigma},\tag{15f}$$

$$\eta Y_o^R \le \sigma \bar{Y}^R,\tag{15g}$$

$$X_o^R \le \sigma \bar{X}^R,\tag{15h}$$

$$\lambda \ge \mathbf{0}, \ \sigma, \eta \ge 0. \tag{15i}$$

All observed DMUs (X_j, Y_j) , $j \in J$, and the bounds on the ratio measures (1) in program (15) are fully scaled by the scaling factor σ . Note that the factor σ is not explicitly present in constraints (15b) and (15c). However, taking into account (15f), the conical combinations of the volume inputs and outputs of the observed DMUs in (15b) and (15c) can also be viewed as the convex combinations of the scaled vectors σY_j^V and σX_j^V taken with the weights λ_j/σ that add up to 1.

7 Returns to scale in the R-VRS technology

If a DMU is scale inefficient, a further question arises: is it too small or too large compared to its MPSS? Answering this question leads to the RTS characterization of the production frontier. For the conventional VRS technology, such characterization is based on the underlying notion of (one-sided) scale elasticity. It was originally developed by Banker (1984) and Banker and Thrall (1992) and further explored in the literature—see, e.g., Førsund and Hjalmarsson (2004), Hadjicostas and Soteriou (2006), Chambers and Färe (2008), Zelenyuk (2013) and Sahoo and Tone (2015).

In this section, we show that the notion of RTS can also be extended to the R-VRS technology. Because the R-VRS technology is generally nonconvex, such extension is not straightforward. In particular, the scale elasticity as a marginal scale characteristic and the RTS characterization based on it are not generally suitable indicators of a direction to MPSS.

In order to extend the notion of RTS to the R-VRS technology, we employ two approaches, depending on the selected sets I' and O' of inputs and outputs with respect to which we measure the scale efficiency. Namely, if the selected inputs and outputs are volume measures, we employ a variant of the standard notion of local RTS (Banker and Thrall 1992). If at least one of the selected inputs or outputs is a ratio measure, we identify a direction to MPSS by exploring the range of optimal values of variable σ in program (13). This approach is conceptually related to the approach of Färe et al. (1983b, 1985) and leads to the notion of global RTS (Podinovski 2004a, b).

The returns-to-scale characterization applies to DMUs located on the production frontier.¹⁰ We therefore require that DMU (X_o, Y_o) satisfies the following assumption:

¹⁰ In the literature, the RTS characterization is sometimes also extended to projections of inefficient DMUs on the frontier, which are not unique and depend on the direction of projection. While such extensions are unproblematic, in order to simplify the exposition, and without loss of generality, we assume that the DMU (X_o, Y_o) under the investigation is technically efficient with respect to the chosen set of outputs O'.

Assumption 4 DMU (X_o, Y_o) is output radial efficient with respect to the subset of outputs O', i.e., in program (5), we have $\tilde{\eta} = 1$.

7.1 Local RTS in the R-VRS technology

Below we define the *local* RTS characterization of the production frontier of technology $T_{\text{VRS}}^{\text{R}}$ with respect to the volume inputs and outputs only, while keeping the ratio inputs and outputs fixed.

Let the sets I' and O' include all volume measures and exclude all ratio measures: $I' = I^V$, $O' = O^V$. Also, let DMU (X_o , Y_o) satisfy Assumption 4. The conventional RTS characterization of the VRS production frontier is determined by the one-sided (left-hand and right-hand) scale elasticities evaluated at DMU (X_o , Y_o). This RTS characterization is local in nature as it depends on the production frontier in a marginal neighborhood of the DMU (X_o , Y_o).

In the case of technology $T_{\text{VRS}}^{\text{R}}$, we can similarly define local RTS (with respect to volume inputs and outputs only) by evaluating the one-sided scale elasticities at the DMU (X_o , Y_o) in the section ΔT of technology $T_{\text{VRS}}^{\text{R}}$ obtained for the fixed vectors X_o^R and Y_o^R . The section ΔT can be regarded as a technology which represents all volume input and output combinations (X^V, Y^V) possible for the fixed vectors X_o^R and Y_o^R of the ratio inputs and outputs.

Taking into account Theorem 1, technology ΔT is the set of all nonnegative DMUs (X^V, Y^V) for which there exists a vector $\lambda \in \mathbb{R}^n$ such that the following conditions are true (note that the vectors X_o^R and Y_o^R are fixed, and there is no need to specify bounds (1) on the ratio measures):

$$\begin{split} &\sum_{j \in J} \lambda_j Y_j^V \geq Y^V, \\ &\sum_{j \in J} \lambda_j X_j^V \leq X^V, \\ &\lambda_j \left(Y_j^R - Y_o^R \right) \geq \mathbf{0}, \qquad \forall j \in J, \\ &\lambda_j \left(X_j^R - X_o^R \right) \leq \mathbf{0}, \qquad \forall j \in J, \\ &\mathbf{1}^\top \lambda = 1, \\ &\lambda \geq \mathbf{0}. \end{split}$$

Technology ΔT can be viewed as the standard VRS technology generated by the subset J' of observed DMUs (X_j, Y_j) such that both vector inequalities $Y_j^R \ge Y_o^R$ and $X_j^R \le X_o^R$ are true. If an observed DMU (X_j, Y_j) does not satisfy these inequalities, we have $\lambda_j = 0$, which excludes this DMU from the subset J'.

By Assumption 4, DMU (X_o, Y_o) is output radial efficient in technology ΔT . Denote $\varepsilon^+(X_o^V, Y_o^V)$ and $\varepsilon^-(X_o^V, Y_o^V)$ the partial right-hand and left-hand scale elasticities evaluated at the DMU (X_o^V, Y_o^V) on the boundary of ΔT . Because ΔT is a VRS technology, their computation is straightforward. Indeed, let ω^{\min} and ω^{\max} be the minimum and maximum optimal values of the dual variable ω to the normalizing equality $\mathbf{1}^\top \lambda = 1$ of the output-oriented linear program stated for DMU (X_o, Y_o) . (Identifying ω^{\min} and ω^{\max} requires solving two linear programs.) Then

$$\varepsilon^+(X_o^V, Y_o^V) = 1 - \omega^{\max}, \quad \varepsilon^-(X_o^V, Y_o^V) = 1 - \omega^{\min}.$$
 (16)

In line with Banker and Thrall (1992), we have the following definition.

Definition 3 DMU $(X_o, Y_o) \in T_{\text{VRS}}^{\text{R}}$ exhibits the following types of *partial* RTS evaluated with respect to the volume inputs and outputs only:

- (i) IRS if $1 < \varepsilon^+(X_o, Y_o) \le \varepsilon^-(X_o, Y_o)$; (ii) DRS if $\varepsilon^+(X_o, Y_o) \le \varepsilon^-(X_o, Y_o) < 1$;
- (iii) CRS if $\varepsilon^+(X_o, Y_o) \le 1 \le \varepsilon^-(X_o, Y_o)$.

Remark 3 The one-sided scale elasticities $\varepsilon^+(X_o^V, Y_o^V)$ and $\varepsilon^-(X_o^V, Y_o^V)$ can also be calculated using the linear program that evaluates the *input* radial efficiency of the DMU (X_o^V, Y_o^V) in the VRS technology ΔT (Førsund and Hjalmarsson 2004; Hadjicostas and Soteriou 2006; Sahoo and Tone 2015; Zelenyuk 2013). Furthermore, in some applications, only a subset of volume inputs and outputs may be of interest and included in these sets, i.e., we may have $I' \subseteq I^V, O' \subseteq O^V$. The partial one-sided scale elasticities with respect to the selected subsets I' and O' can be evaluated in technology ΔT using the linear programming approach of Podinovski et al. (2016) of which formulae (16) are a special case.

7.2 Global RTS in the R-VRS technology

The *global* RTS (GRS) characterization of production frontiers was introduced by Podinovski (2004a, b). The types of GRS are indicative of the direction in which DMU (X_o , Y_o) should resize to achieve MPSS. For example, DMU (X_o , Y_o) exhibits global increasing RTS if it is smaller than its MPSS and, therefore, needs to increase the scale of its operations to achieve its MPSS.¹¹

In any convex technology, including the VRS technology, the local and global RTS characterizations are identical (Podinovski 2017). However, in a nonconvex technology such as the R-VRS technology, the local types of RTS are no longer indicative of the direction to MPSS (see the example in Sect. 7.3), and the global characterization needs to be used instead.

In this section, we consider arbitrary nonempty subsets of volume and ratio inputs and outputs $I' \subseteq I$ and $O' \subseteq O$. Let DMU (X_o, Y_o) satisfy Assumption 4. In order to evaluate its scale efficiency with respect to the selected sets I' and O', we solve program (9) or equivalent program (13).

Any optimal solution $\langle \varphi^*, \psi^* \rangle$ of program (9) defines the MPSS of DMU (X_o, Y_o) evaluated with respect to the inputs and outputs from the sets I' and O' and stated as

$$(X_o(\varphi^*), Y_o(\psi^*)) \in T_{\text{VRS}}^{\text{R}}.$$
(17)

Assume that DMU (X_o , Y_o) is scale *inefficient*, i.e., that $\psi^*/\varphi^* > 1$. Then either $\varphi^* < 1$ or $\varphi^* > 1$.¹² If $\varphi^* < 1$, the corresponding MPSS (17) is smaller than DMU (X_o , Y_o). If $\varphi^* > 1$, the MPSS (17) is larger than DMU (X_o , Y_o). It is theoretically possible that program (9) has alternative optimal solutions (φ^*, ψ^*), each identifying a different MPSS (17).

By statement (iii) of Theorem 5, the case $\varphi^* < 1$ corresponds to $\hat{\sigma} > 1$ in an optimal solution of program (13), and the case $\varphi^* > 1$ corresponds to $\hat{\sigma} < 1$. We can now extend the

¹¹ Mostafaee and Soleimani-damaneh (2020) suggest a further refinement of the GRS characterization by identifying two subtypes of each of the G-IRS and D-GRS types of global RTS.

¹² The case $\varphi^* = 1$ is impossible because, by Assumption 4, if $\varphi^* = 1$ then $\psi^* = 1$ and $\psi^* / \varphi^* = 1$, which contradicts the assumption that DMU (X_o , Y_o) is scale inefficient. Also, $\varphi^* > 1$ implies $\psi^* > 1$, and $\varphi^* < 1$ implies $\psi^* \leq 1$. If, in addition to Assumption 4, DMU (X_o , Y_o) is input radial efficient with respect to the subset of inputs I', then $\varphi^* < 1$ implies $\psi^* < 1$ (Podinovski 2004a).

characterization of GRS defined in Podinovski (2004a) in the case I' = I and O' = O to the case of *partial* GRS evaluated only with respect to the sets of inputs and outputs I' and O'.

Let DMU $(X_o, Y_o) \in T_{\text{VRS}}^{\text{R}}$ satisfy Assumption 4. In order to determine if DMU (X_o, Y_o) is at MPSS and, if not, whether there exists an MPSS that is smaller or larger than DMU (X_o, Y_o) , we solve two additional programs derived from program (13). The first can be viewed as the non-increasing returns-to-scale (NIRS) analogue of program (13), and the second as its non-decreasing returns-to-scale (NDRS) analogue.

$$\hat{\eta}_1 = \max \eta$$

subject to $(13b) - (13c)$ and $\sigma < 1$. (18)

and

$$\hat{\eta}_2 = \max \eta$$

subject to (13b) - (13o) and $\sigma > 1$. (19)

It is clear that $\hat{\eta}_1 \ge 1$ and $\hat{\eta}_2 \ge 1$. We also have $\hat{\eta} = \min\{\hat{\eta}_1, \hat{\eta}_2\}$, where $\hat{\eta}$ is the optimal value of program (13). The following four mutually exclusive cases are now possible.

If $\hat{\eta}_1 = \hat{\eta}_2 = 1$, the optimal value $\hat{\eta}$ of program (13) is equal to 1 and DMU (X_o, Y_o) is at MPSS with respect to the sets I' and O'. In this case, we class DMU (X_o, Y_o) as exhibiting global CRS (G-CRS).

Let $1 \le \hat{\eta}_2 < \hat{\eta}_1$. Then $\hat{\eta} = \hat{\eta}_1 > 1$ and, for any optimal solution $\langle \hat{\lambda}, \hat{\sigma}, \hat{\eta} \rangle$ to (13), we have $\hat{\sigma} < 1$. By statement (iii) of Theorem 5, for any optimal solution $\langle \psi^*, \varphi^* \rangle$ of program (9), we have $\varphi^* = 1/\hat{\sigma} > 1$. The MPSS of DMU (X_o, Y_o) calculated by formula (17) is larger than DMU (X_o, Y_o), for any optimal solution of (13), and we class DMU (X_o, Y_o) as exhibiting global IRS (G-IRS).

Similarly, if $1 \le \hat{\eta}_1 < \hat{\eta}_2$, then for any optimal solution $\langle \hat{\lambda}, \hat{\sigma}, \hat{\eta} \rangle$ to (13) we have $\hat{\sigma} > 1$ and, in program (9), $\varphi^* = 1/\hat{\sigma} < 1$. Therefore, any MPSS of DMU (X_o, Y_o) is smaller than this DMU and the latter exhibits global DRS (G-DRS).

The final logical possibility is the case in which $1 < \hat{\eta}_1 = \hat{\eta}_2$. In this case, DMU (X_o, Y_o) is scale inefficient and has at least two different directions to MPSS, one of which is larger, and the other smaller, than the DMU (X_o, Y_o) . We class DMU (X_o, Y_o) as exhibiting global sub-constant RTS (G-SCRS).

Remark 4 Assume that DMU (X_o, Y_o) does not satisfy Assumption 4, i.e., it is not efficient with respect to the vector of outputs included in the set O'. It is straightforward to show that, in this case, the described approach will characterize GRS at the projection of DMU (X_o, Y_o) on the boundary of technology $T_{\text{VRS}}^{\text{R}}$. Indeed, let the efficiency of DMU (X_o, Y_o) in technology $T_{\text{VRS}}^{\text{R}}$ evaluated with respect to the outputs in the set O' by program (5) be equal to $1/\tilde{\eta} < 1$. Denote $(X^*, Y^*) = (X_o, Y_o(\tilde{\eta}))$ the projection of DMU (X_o, Y_o) on the boundary of technology $T_{\text{VRS}}^{\text{R}}$. Then the projected DMU (X^*, Y^*) satisfies Assumption 4 and we can characterize its type of GRS by solving programs (18) and (19). Let the corresponding optimal values of these two programs be $\hat{\eta}_1^*$ and $\hat{\eta}_2^*$, respectively. It is clear that $\hat{\eta}_1^* = \hat{\eta}_1/\tilde{\eta}$ and $\hat{\eta}_2^* = \hat{\eta}_2/\tilde{\eta}$, where $\hat{\eta}_1$ and $\hat{\eta}_2$ are the optimal values of programs (18) and (19) stated for the evaluation of DMU (X_o, Y_o) . Then $\hat{\eta}_1^* \leq \hat{\eta}_2^*$ if and only if $\hat{\eta}_1 \leq \hat{\eta}_2$. This implies that the procedure for the characterization of GRS based on the optimal values $\hat{\eta}_1$ and $\hat{\eta}_2$ of programs (18) and (19) stated for DMU (X_o, Y_o) (at which the notion of GRS is undefined) in fact characterizes GRS at the projection (X^*, Y^*) .

7.3 Example of evaluation of RTS

Let $T_{\text{VRS}}^{\text{R}}$ be the R-VRS technology considered in Examples 2 and 3. We first illustrate the notion of local RTS, by referring to scenario (ii) in which we define I' = I and O' = O. In this case, technology ΔT is the section *KCADE* shown in Fig. 1 and, separately, in Fig. 6.

Consider DMU *A* and note that it satisfies Assumption 4. The left-hand scale elasticity at DMU *A* is undefined because it is not possible to reduce its input in technology ΔT , and we can conventionally take $\varepsilon^{-}(A) = +\infty$. The right-hand scale elasticity $\varepsilon^{+}(A) = 2$. It corresponds to the movement along the side *AD* away from *A* and can be calculated as the ratio of the marginal productivity (equal to the slope 2 of the line *AD*) to the average productivity at *A* (equal to 1/1 = 1). By Definition 3, DMU *A* exhibits IRS with respect to the volume input and output.

Because the section ΔT is convex, the local type of RTS of DMU A coincides with its global (G-IRS) type, and both are indicative of the direction to its MPSS at DMU D.

Now consider scenario (i) in which the set I' includes the single volume input and the set O' includes both volume and ratio outputs. As shown in Example 3, DMU A is scale inefficient and its MPSS is DMU P. Because DMU P is larger than A, the latter exhibits G-IRS with respect to all inputs and outputs. Note that, in this scenario, the right-hand scale elasticity $\varepsilon^{-}(A) = 0$ and corresponds to the movement along the line AL (see Footnote 5). Therefore, locally, DMU A exhibits DRS to the right, although in the global sense, it should increase the scale of its operations to achieve its MPSS at DMU P.

Finally, consider scenario (iii) in which the set O' includes only the ratio output. As shown in Example 3, DMU A is scale inefficient and its MPSS is at DMU V, which is larger than A. Therefore, DMU A exhibits G-IRS in this scenario as well. As in scenario (i), the right-hand scale elasticity at A in the section of technology $T_{\text{VRS}}^{\text{R}}$ defined by its boundary ALVW is equal to zero and corresponds to the marginal movement along the line AL away from A. Therefore, DMU A exhibits DRS, and this local characterization is not indicative of the direction to MPSS.

8 Illustrative application

In this section, we illustrate the application of the developed methodology using a sample of 39 secondary schools in the West Midlands region of England. The data was collected in 2020-21 and is publicly available from the official website of Department for Education.





Measure	Mean	Median	Minimum	Maximum	Standard deviation	
Input 1: teachers	66.12	63.9	22.2	124.9	22.19	
Input 2: expenses	2773.69	2543.9	1184	5638.9	1127.32	
Input 3: NFSM	79.33	81.3	55.7	96.8	11.31	
Output 1: pupils	1070.08	1059	396	19.33	344.73	
Output 2: HE	0.69	0.72	0.45	0.91	0.12	

 Table 5
 Descriptive statistics for the application

Table 5 shows summary statistics of the measures used in the application. (The full data set used in this application is given in Appendix C.) The volume inputs 1 and 2 represent the number of teachers and expenses in thousands of British pounds (excluding teacher salaries) at each school. The ratio input 3 measures the percentage of pupils who are not eligible for free school meals (NFSM). This input is a socio-economic characteristic of the catchment area which is assumed to have a positive effect on the performance of the school. The volume output 1 is the number of all pupils at the school. The ratio output 2 is the proportion of the final year (sixth form) pupils proceeding to higher education (HE) and is included in the model as a measure of quality of education. (In Sect. 8.4, we discuss potential computational issues arising from the use of small ratios in DEA models.¹³)

We consider three scenarios of evaluation of scale characteristics, by choosing the sets I' and O' in different ways: first, as the sets of all (volume and ratio) measures; second, as the sets of all volume inputs and outputs; and third, as the sets of all ratio inputs and outputs, respectively.

Preliminary calculations show that 23 schools in the current sample are strongly (Pareto) efficient and, therefore, satisfy Assumption 4. It also turns out that, in each scenario, no additional observed DMUs satisfy this assumption, i.e., the set of observed DMUs that are efficient with respect to the outputs in the set O' is the set of strongly efficient DMUs of technology $T_{\text{VRS}}^{\text{R}}$.¹⁴

8.1 Evaluation with respect to both volume and ratio measures

Consider the evaluation of scale efficiency and RTS characterization of the schools with respect to all, volume and ratio, inputs and outputs. In this scenario, we define the sets I' = I and O' = O.

Table 6 shows the results of computations. Its second column shows the output radial efficiency of each school in the R-VRS technology $T_{\text{VRS}}^{\text{R}}$ evaluated with respect to the vector of two outputs. The next three columns show the output radial efficiency of the schools in models (13), (18) and (19) which are based on the closed full cone extension $\bar{C}(I, O)$ of

 $^{^{13}}$ We use percentages for input 3 and fractions between 0 and 1 for output 2 in line with the way the data are reported by Department for Education. An obvious alternative would be to convert percentages to fractions or fractions to percentages. Rescaling data in such a way does not affect radial measures of efficiency (evaluated with respect to the set O') and scale characteristics based on them.

¹⁴ To verify strong efficiency of DMU (X_o , Y_o), we solve the additive model that seeks the maximum sum of individual improvements (slacks in the input and output constraints) of all volume and ratio inputs and outputs, under the condition that the resulting DMU is in technology $T_{\text{VRS}}^{\text{R}}$. This model is stated as model (12) in Olesen et al. (2017) in which we take $(\hat{X}, \hat{Y}) = (X_o, Y_o)$.

School	R-VRS	Cone model (13)	NIRS model (18)	cone	NDRS model (19)	cone	Scale effi- ciency	Global RTS
1	0.9337	0.9122	0.9122		0.9147		0.977	G-IRS
2	1	1	1		1		1	MPSS
3	1	1	1		1		1	MPSS
4	1	1	1		1		1	MPSS
5	1	1	1		1		1	MPSS
6	0.8462	0.8163	0.8462		0.8163		0.9647	G-DRS
7	1	1	1		1		1	MPSS
8	1	1	1		1		1	MPSS
9	1	1	1		1		1	MPSS
10	1	1	1		1		1	MPSS
11	1	0.9463	0.9463		1		0.9463	G-IRS
12	0.7066	0.6974	0.7066		0.6974		0.987	G-DRS
13	0.9753	0.9391	0.9664		0.9391		0.9629	G-DRS
14	1	1	1		1		1	MPSS
15	1	0.9779	0.9779		1		0.9779	G-IRS
16	1	0.8618	0.8618		1		0.8618	G-IRS
17	1	0.9898	0.9898		1		0.9898	G-IRS
18	0.8853	0.8704	0.8853		0.8704		0.9832	G-DRS
19	0.9881	0.9308	0.9789		0.9308		0.942	G-DRS
20	0.9427	0.9292	0.9292		0.9322		0.9857	G-IRS
21	1	0.8955	0.8955		1		0.8955	G-IRS
22	0.8941	0.7919	0.8941		0.7919		0.8857	G-DRS
23	0.9733	0.9352	0.9733		0.9352		0.9609	G-DRS
24	1	0.9813	1		0.9813		0.9813	G-DRS
25	0.9765	0.8477	0.9765		0.8477		0.8681	G-DRS
26	0.8941	0.8908	0.8941		0.8908		0.9963	G-DRS
27	1	0.9105	0.9105		1		0.9105	G-IRS
28	0.9519	0.9225	0.9519		0.9225		0.9691	G-DRS
29	0.8927	0.8706	0.8706		0.8757		0.9752	G-IRS
30	1	1	1		1		1	MPSS
31	1	0.7711	0.7711		1		0.7711	G-IRS
32	1	0.8869	0.8869		1		0.8869	G-IRS
33	1	1	1		1		1	MPSS
34	0.8012	0.7712	0.8012		0.7712		0.9626	G-DRS
35	1	1	1		1		1	MPSS
36	0.9017	0.8767	0.8834		0.8767		0.9723	G-DRS
37	0.7406	0.7273	0.7406		0.7273		0.982	G-DRS
38	1	1	1		1		1	MPSS
39	1	1	1		1		1	MPSS

 Table 6
 Efficiency and scale characteristics evaluated with respect to both volume and ratio measures (for inefficient schools, see Remark 5)

technology $T_{\text{VRS}}^{\text{R}}$ and its NIRS and NDRS analogues. Note that the efficiencies shown in Table 6 are inverse to the optimal values $\hat{\eta}$, $\hat{\eta}_1$ and $\hat{\eta}_2$ of these programs.¹⁵

Following the discussion of Sect. 6.3, instead of solving program (13), we solve the simpler, but equivalent, program (15). We also solve this program with the additional constraints $\sigma \le 1$ and $\sigma \ge 1$, instead of the NIRS and NDRS programs (18) and (19).

According to Sect. 4.4, the overall scale efficiency of each school is equal to its efficiency in model (13), as shown in the third column of Table 6. The scale efficiency of each school is shown in the second last column of this table. It is obtained as the ratio of its output efficiency in model (13) to its efficiency in technology $T_{\text{VRS}}^{\text{R}}$.

Let us now consider the RTS characterization of schools located on the frontier of technology $T_{\text{VRS}}^{\text{R}}$. As highlighted in Sect. 7, because this technology is not convex, the standard RTS characterization of its frontier is generally not well-defined and is not indicative of the direction to MPSS. Instead, we use the global RTS characterization to identify the direction to MPSS for each school.

As already established, 23 schools from the sample are output radial efficient and are located on the frontier of technology $T_{\text{VRS}}^{\text{R}}$. (The efficiency of these schools in the R-VRS model is equal to 1, as shown in the second column of Table 6.) As discussed in Sect. 7.2, the characterization of GRS for these schools can be obtained by comparing their efficiency in the NIRS and NDRS programs (18) and (19).

The last column of Table 6 shows the resulting GRS characterization. For example, School 2 is at MPSS (or, equivalently, exhibits G-CRS). School 11 exhibits G-IRS and is, therefore, smaller than its MPSS. (Note that, in this case, for the optimal values of programs (18) and (19), we have $\hat{\eta}_1 = 1/0.9463 > 1/1 = \hat{\eta}_2$). School 24 exhibits G-DRS and is, therefore, larger than its MPSS.

Remark 5 For the inefficient schools, the last column of Table 6 shows the GRS characterization of their radial projections on the boundary of technology $T_{\text{VRS}}^{\text{R}}$ (see Remark 4 for the discussion of GRS evaluated at the projections of inefficient DMUs). For example, School 1 is inefficient. Improving its efficiency by proportional increase of both outputs (by the factor 1/0.9337 = 1.071) would result in the school that exhibits G-IRS. Such projection would be smaller than its MPSS.

8.2 Evaluation with respect to volume measures only

Consider the evaluation of scale efficiency and RTS with respect to volume measures only, by keeping the ratio input and output fixed. Define $I' = \{input 1, input 2\}$ and $O' = \{output 1\}$. In this scenario, we investigate the relationship between the volume parameters of the schools (teachers, expenses and pupils) by assuming that the socio-economic characteristic of the school catchment area (ratio input 3) and quality of education (ratio output 2) remain unchanged.

In this case, model (13) becomes program (14). Models (18) and (19) are obtained from the latter program by the reinstatement of constraint (13j) and the incorporation of additional constraints $\sigma \leq 1$ and $\sigma \geq 1$, respectively.

¹⁵ Solving the output-oriented R-VRS model as a mixed integer linear program was discussed in Sect. 3.2 of Olesen et al. (2017). Solving the cone model (13) and its NIRS and NDRS variants (18) and (19) (in all three scenarios considered in the current application) is similar to solving the R-CRS models, as also discussed by Olesen et al. (2017). In all these models, we restate the nonlinear inequalities (13f) and (13h) as "either-or" conditions and linearize them using the common "Big *M*" approach. The big constant *M* can be assessed using the constraints of program (13). In all our computations, we used the same value M = 100.

Table 7 shows the results of computations in a format similar to Table 6. The difference with the previous scenario is that the efficiencies of all schools are now evaluated with respect to the volume output 1 only.¹⁶ Also, as discussed in Sect. 7.1, in this scenario, the local and global characterizations of RTS are identical.

Consider for example, School 3. As shown in Table 7, this school exhibits DRS (or G-DRS in the equivalent terminology of GRS). This means that a marginal proportional increase of the number of teachers and expenses of this school would result in a smaller rate of increase of the number of pupils, assuming that the ratio measures (input 3 and output 2) remain constant. From a policy-maker's perspective, School 3 is scale inefficient and moving to MPSS means making the school smaller, without any change to the socio-economic environment and without affecting the quality of education.

8.3 Evaluation with respect to ratio measures only

Suppose that we are interested in the impact of the socio-economic environment on the proportion of pupils going to university, while keeping the volume measures fixed. Define $I' = \{\text{input } 3\}$ and $O' = \{\text{output } 2\}$.

Table 8 shows the results of computations using appropriately specified programs (13), (18) and (19).

For example, consider School 11. Because it exhibits G-IRS, it is smaller than its MPSS evaluated with respect to the ratio input and output. This implies that, if we keep the size of the school unchanged (represented by the number of teachers, expenses and pupils), then the socio-economic characteristic of the pupil intake (percentage of pupils not eligible for free school meals) has a more than proportional impact on the quality of education (proportion of pupils going to university).

8.4 Notes on computations with small ratios

As mentioned in Footnote 13, in the discussed application, output 2 represents the proportion of final-year pupils proceeding to higher education. We could, of course, convert these proportions to percentages, by rescaling the proportions by a factor of 100. However, we preferred to keep this output unchanged, in line with the way the data was reported by Department for Education.

Using proportions and any other very small numerical values (compared, for example, with the numerical values for school expenditure all of which are in excess of 1000) in DEA models may present well-known computational problems, and rescaling the data prior to calculations may be required in order to obtain correct results. (In the reported application, we were aware of potential problems and repeated our computations with the rescaled ratios, but did not detect any discrepancy between the results.)

The scaling of linear optimization problems is a technique utilized in linear optimization solvers with the aim to improve the conditioning of the constraint matrix and decrease the computational effort for solution—see, e.g., Bixby (2002). As observed by Elble and Sahini-dis (2012), scaling provides a relative point of reference for absolute tolerances. This is especially important when solving the linearized counterparts of models (13), (18) and (19),

¹⁶ Also, in this case, the cone extension of the R-VRS technology and its NIRS and NDRS variants are obtained with respect to the volume measures only. These sets are different from the cone (and NIRS and NDRS) extensions obtained with respect to all, volume and ratio, measures. Therefore, the efficiency scores in Tables 6 and 7, and also in Table 8, are not directly comparable.

	-			1		·
School	R-VRS	Cone model (13)	NIRS cone model (18)	NDRS cone model (19)	Scale effi- ciency	Global (=local) RTS
1	0.9239	0.9041	0.9239	0.9041	0.9786	DRS
2	1	1	1	1	1	MPSS
3	1	0.9453	1	0.9453	0.9453	DRS
4	1	1	1	1	1	MPSS
5	1	1	1	1	1	MPSS
6	0.7773	0.6746	0.7773	0.6746	0.8679	DRS
7	1	1	1	1	1	MPSS
8	1	0.9765	1	0.9765	0.9765	DRS
9	1	0.9408	1	0.9408	0.9408	DRS
10	1	1	1	1	1	MPSS
11	1	0.9622	0.9622	1	0.9622	IRS
12	0.7066	0.6228	0.7066	0.6228	0.8814	DRS
13	0.9753	0.7754	0.9753	0.7754	0.795	DRS
14	1	0.9111	1	0.9111	0.9111	DRS
15	1	1	1	1	1	MPSS
16	1	0.9935	1	0.9935	0.9935	DRS
17	1	0.8534	1	0.8534	0.8534	DRS
18	0.8853	0.8553	0.8853	0.8553	0.9661	DRS
19	0.9377	0.9279	0.9377	0.9279	0.9895	DRS
20	0.9427	0.9322	0.9427	0.9322	0.9889	DRS
21	1	0.811	0.811	1	0.811	IRS
22	0.846	0.6793	0.846	0.6793	0.803	DRS
23	0.9518	0.8707	0.9518	0.8707	0.9148	DRS
24	1	0.7734	1	0.7734	0.7734	DRS
25	0.8874	0.7504	0.8874	0.7504	0.8456	DRS
26	0.7238	0.6749	0.7238	0.6749	0.9324	DRS
27	1	1	1	1	1	MPSS
28	0.9519	0.9225	0.9519	0.9225	0.9691	DRS
29	0.8927	0.7875	0.8927	0.7875	0.8822	DRS
30	1	0.8661	1	0.8661	0.8661	DRS
31	1	1	1	1	1	MPSS
32	1	0.9019	0.9019	1	0.9019	IRS
33	1	1	1	1	1	MPSS
34	0.7876	0.7543	0.7876	0.7543	0.9577	DRS
35	1	1	1	1	1	MPSS
36	0.9017	0.8441	0.9017	0.8441	0.9361	DRS
37	0.7339	0.6653	0.7339	0.6653	0.9065	DRS
38	1	1	1	1	1	MPSS
39	1	0.8866	0.8866	1	0.8866	IRS

Table 7 Efficiency and scale characteristics evaluated with respect to volume measures only

School	R-VRS	Cone model (13)	NIRS cone model (18)	NDRS cone model (19)	Scale ciency	effi-	Global RTS
1	0.8276	0.7662	0.8276	0.7662	0.92581		G-DRS
2	1	1	1	1	1		MPSS
3	1	1	1	1	1		MPSS
4	1	1	1	1	1		MPSS
5	1	1	1	1	1		MPSS
6	0.7912	0.5885	0.7912	0.5885	0.743807		G-DRS
7	1	0.9408	0.9408	1	0.9408		G-IRS
8	1	1	1	1	1		MPSS
9	1	0.9179	0.9179	1	0.9179		G-IRS
10	1	1	1	1	1		MPSS
11	1	0.8686	0.8686	1	0.8686		G-IRS
12	0.6322	0.4562	0.6322	0.4562	0.721607		G-DRS
13	0.8706	0.8054	0.8706	0.8054	0.925109		G-DRS
14	1	0.95	0.95	1	0.95		G-IRS
15	1	1	1	1	1		MPSS
16	1	0.6625	0.6625	1	0.6625		G-IRS
17	1	0.8642	0.8642	1	0.8642		G-IRS
18	0.5952	0.5736	0.5952	0.5736	0.96371		G-DRS
19	0.9881	0.9006	0.9789	0.9006	0.911446		G-DRS
20	0.8391	0.7222	0.8391	0.7222	0.860684		G-DRS
21	1	0.9031	0.9031	1	0.9031		G-IRS
22	0.8941	0.7583	0.8941	0.7583	0.848115		G-DRS
23	0.9733	0.9482	0.9733	0.9482	0.974211		G-DRS
24	1	1	1	1	1		MPSS
25	0.9765	0.8342	0.9765	0.8342	0.854275		G-DRS
26	0.8941	0.8196	0.8941	0.8196	0.916676		G-DRS
27	1	1	1	1	1		MPSS
28	0.7059	0.6343	0.7059	0.6343	0.898569		G-DRS
29	0.6471	0.5892	0.6471	0.5892	0.910524		G-DRS
30	1	1	1	1	1		MPSS
31	1	1	1	1	1		MPSS
32	1	0.8869	0.8869	1	0.8869		G-IRS
33	1	1	1	1	1		MPSS
34	0.7176	0.6707	0.7176	0.6707	0.934643		G-DRS
35	1	1	1	1	1		MPSS
36	0.5647	0.5345	0.5647	0.5345	0.94652		G-DRS
37	0.5977	0.5048	0.5977	0.5048	0.844571		G-DRS
38	1	1	1	1	1		MPSS
39	1	1	1	1	1		MPSS

 Table 8
 Efficiency and scale characteristics evaluated with respect to ratio measures only

because the linearization of constraints (13f) and (13h) involves one binary variable used with one "Big M" for each DMU in the sample. (The value of "Big M" should of course be chosen sufficiently large but not much larger than necessary. To avoid too large "Big M" values, it is of course valid to use two different "Big M" values, one for constraints (13f) and the other for (13h).)

Absolute tolerances are applied to every node of the branch-and bound tree used in the simplex method to determine when a binary variable in a relaxed solution is considered to be integer or a reduced cost coefficient is considered to be nonnegative. The use of absolute tolerances is challenging when, for example, rows in the constraint matrix are of different numerical magnitude as in the constraint matrix for the data set in Table 10. For example, in the discussed application, the numerical values for school expenditure are up to five orders of magnitude larger than the corresponding proportions of pupils proceeding to higher education. A further computational challenge may arise with the use of the "Big *M*" approach to the linearization of optimization programs with an unbalanced constraint matrix.

In the reported application, we could balance the constraint matrix by restating expenditures in £100,000, pupils in hundreds and their proportion going to higher education as percentages. This would result in all data being in the range [0; 100]. It has been decided to maintain data as is rather than to rescale it. This proved not to be a problem for industry-leading solvers, e.g., CPLEX or Gurobi, because their presolvers include sophisticated rescaling of the constraint matrix by default.

9 Conclusion

The notions of the most productive scale size, scale and overall efficiency play an important role in efficiency and productivity analysis based on the conventional VRS model. This paper develops an extension of such scale characteristics to the R-VRS technology of Olesen et al. (2015). The latter technology is defined axiomatically and allows both volume and ratio inputs and outputs to be incorporated in the model.

Following the approach of Banker (1984), we start by defining the notion of MPSS in the R-VRS technology. Continuing with the same approach, we define and interpret scale efficiency as a measure of divergence from MPSS and decompose the overall efficiency of a DMU into its technical and scale components.

Similar to the case of conventional VRS technology, the evaluation of MPSS and scale efficiency of a DMU in the R-VRS technology turns out to be equivalent to the assessment of its output radial efficiency in the closed cone extension of the R-VRS technology. Obtaining an explicit statement of such cone extension suitable for optimization is a nontrivial task accomplished in our paper.

Despite the conceptual similarities between the cases of VRS and R-VRS technologies, there is an important distinction to be noted. In the former case, the closed cone extension of the VRS technology is the CRS technology. This allows the scale efficiency in the VRS technology to be alternatively interpreted as the technical efficiency in the benchmark CRS technology. However, as shown in this paper, the closed cone extension of the R-VRS technology is generally different from the R-CRS technology axiomatically defined by Olesen et al. (2015). As demonstrated by examples in this paper, the MPSS and scale efficiency in the R-VRS technology defined according to the approach of Banker (1984) cannot generally be evaluated using the R-CRS technology.

In order to keep the exposition more general and suitable for different applications, we allow the evaluation of MPSS and scale efficiency with respect to any selected subsets of inputs and outputs, while keeping the remaining measures constant. This corresponds to the scenario in which we want to test the response of a selected subset of outputs to changes of a selected subset of inputs while treating the other inputs and outputs as fixed exogenous measures.

We also consider returns-to-scale characterizations of the production frontier of the R-VRS technology. Depending on the choice of inputs and outputs with respect to which we evaluate the MPSS and scale efficiency of a DMU, we employ the local and global variants of the notion of returns to scale. For a scale inefficient DMU, the type of returns to scale (increasing or decreasing) is indicative of the direction to its MPSS, when the resizing is allowed only with respect to the selected inputs and outputs.

To outline further potential research avenues, it is worth noting that the use of ratio data presents similar difficulties in many alternative convex DEA models, for example the models based on assumptions of weak disposability, models incorporating value judgements in the form of weight restrictions and models with complex network structures. It is clear that ratio inputs and outputs are inconsistent with the assumption of convexity and the additional assumption of scalability often made in such models. Exploring the ways to incorporate ratio data and developing approaches to the efficiency and RTS evaluation in such models is a challenging task for future research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10479-022-04862-6.

Funding Not applicable.

Code availability Supplementary GAMS code is available online.

Declarations

Conflict of interest The authors have no conflicting interests.

Ethical approval Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix A: Proofs

Lemma 1 Define W as the set of all DMUs $(X, Y) \in \mathbb{R}^{m+s}_+$ that satisfy conditions (8) with some λ and σ . Then W is a closed set.

Proof of Lemma 1 Let a sequence of DMUs $\{(X^k, Y^k)\} \subset W$ converge to $(X^*, Y^*) \in \mathbb{R}^{m+s}$ as $k \to +\infty$. We need to prove that $(X^*, Y^*) \in W$.

For each $k = 1, 2, ..., (X^k, Y^k)$ satisfies (8) with some vector λ^k and scalar σ^k . Because the sequence $\{\lambda^k\}$ is contained in the compact set $[0, 1]^n$, where *n* is the number of observed

DMUs, there exists some (not necessarily unique) convergent subsequence $\{\lambda^{k_t}\}, t = 1, 2, ...,$ of the sequence $\{\lambda^k\}$. Let the limit point of $\{\lambda^{k_t}\}$ be λ^* . Because $[0, 1]^n$ is a closed set, $\lambda^* \in [0, 1]^n$. Therefore, $\{\lambda^{k_t}\} \rightarrow \lambda^*$ as $t \rightarrow +\infty$. To simplify the proof, and without loss of generality, let us assume that the sequence $\{\lambda^k\}$ itself is converging, i.e., that $\{\lambda^k\} \rightarrow \lambda^*$ as $k \rightarrow +\infty$.

Let us prove that the sequence $\{\sigma^k\}$ is contained in a compact set. Because $\{X^k\}$ converges to X^* , it is bounded and there exists an M > 0 such that $X_i^k < M$, for all $i \in I$ and k = 1, 2, ...Because λ^* satisfies (8i), there exists a $j' \in J$ such that $\lambda_{j'}^* > 0$. Because $\{\lambda_{j'}^k\} \to \lambda_{j'}^*$ as $k \to +\infty$, there exists an $\varepsilon > 0$ and $\bar{k} \ge 1$ such that $\lambda_{j'}^k \ge \varepsilon > 0$, $\forall k > \bar{k}$.

The inequalities (8c) and (8g) stated for all DMUs $(X^k, Y^k), k > \overline{k}$, imply

$$\lambda_{j'}^k \sigma^k X_{j'i}^V \le (X_i^k)^V, \quad \forall i \in \mathbf{I}^V \cap \mathbf{I}', \lambda_{j'}^k (\sigma^k X_{j'i}^R - (X_i^k)^R) \le 0, \quad \forall i \in \mathbf{I}^R \cap \mathbf{I}'.$$

$$(20)$$

By definition, we have $0 < \varepsilon \le \lambda_{j'}^k$ and $M > X_i^k$, $\forall k > \bar{k}$ and $\forall i \in I$. Then, from the inequalities (20), for all $k > \bar{k}$, we have

$$\varepsilon \sigma^k X^V_{j'i} \le \lambda^k_{j'} \sigma^k X^V_{j'i} \le (X^k_i)^V \le M, \quad \forall i \in \mathbf{I}^V \cap \mathbf{I}',$$
(21a)

$$\sigma^k X_{j'i}^R \le (X_i^k)^R \le M, \qquad \forall i \in \mathbf{I}^R \cap \mathbf{I}'.$$
(21b)

By Assumption 1, there exists an $i' \in I'$ such that $X_{j'i'} > 0$. If $i' \in I^V$, then $X_{j'i'}^V > 0$ and (21a) implies that $\sigma^k \leq \sigma_1$, for all $k > \bar{k}$, where $\sigma_1 = M/(\varepsilon X_{j'i'}^V) > 0$. If $i' \in I^R$, then $X_{j'i'}^R > 0$ and (21b) implies that $\sigma^k \leq \sigma_2$, for all $k > \bar{k}$, where $\sigma_2 = M/X_{j'i'}^R > 0$. Denote $\bar{\sigma} = \min\{\sigma_1, \sigma_2\}$. We have proved that the sequence $\{\sigma^k\}, k \geq \bar{k}$, is contained in the compact set $[0, \bar{\sigma}]$. Then there exists its subsequence that converges to some $\sigma^* \in [0, \bar{\sigma}]$. Without loss of generality, assume that the sequence $\{\sigma^k\}$ itself is converging, i.e., that $\{\sigma^k\} \to \sigma^* \in [0, \bar{\sigma}]$ as $k \to +\infty$.

By assumption, conditions (8) are true for each DMU (X^k, Y^k) with the corresponding λ^k and σ^k , k = 1, 2, ... We also have $(X^k, Y^k) \rightarrow (X^*, Y^*)$, as $k \rightarrow +\infty$. Furthermore, as proved, without loss of generality, we can assume that $\lambda^k \rightarrow \lambda^*$ and $\sigma^k \rightarrow \sigma^*$ as $k \rightarrow +\infty$. This implies that conditions (8) are also true at the limit, i.e., for (X^*, Y^*) , λ^* and σ^* . Therefore, $(X^*, Y^*) \in W$.

Proof of Theorem 2 Define W as in Lemma 1. We need to prove that $\overline{C}(I', O') = W$. First, consider any DMU $(X, Y) = (\tilde{X}(\alpha), \tilde{Y}(\alpha)) \in C(I', O')$, where $(\tilde{X}, \tilde{Y}) \in T_{\text{VRS}}^{\text{R}}$ and $\alpha \ge 0$. DMU (\tilde{X}, \tilde{Y}) satisfies (2) with some vector λ . Then (X, Y) satisfies (8) with the same vector λ and $\sigma = \alpha$. Therefore, $(X, Y) \in W$ and $C(I', O') \subseteq W$. Taking the closure of both sides and noting equality (7) and Lemma 1, we have $\overline{C}(I', O') \subseteq W$.

It remains to be proved that $W \subseteq \overline{C}(I', O')$. Consider any DMU $(\hat{X}, \hat{Y}) \in W$. Then (\hat{X}, \hat{Y}) satisfies (8) with some $\hat{\lambda}$ and $\hat{\sigma}$. We need to prove that $(\hat{X}, \hat{Y}) \in \overline{C}(I', O')$.

The idea of the proof is to consider two possibilities: $\hat{\sigma} > 0$ and $\hat{\sigma} = 0$. In the former case, the proof is straightforward. In the latter case, we first define a sequence of DMUs $\{(X_{(k)}, Y_{(k)})\}, k = 1, 2, ...,$ such that $(X_{(k)}, Y_{(k)}) \in T_{\text{VRS}}^{\text{R}} \subset C(I', O')$. In other words, $(X_{(k)}, Y_{(k)})$ satisfies conditions (2), $\forall k$. In addition, the sequence $\{(X_{(k)}, Y_{(k)})\}$ is constructed in such a way that the DMU (\hat{X}, \hat{Y}) is its limit point. Therefore, $(\hat{X}, \hat{Y}) \in \overline{C}(I', O')$, which completes the proof.

(i) Let $\hat{\sigma} > 0$. Using notation (4), where $\varphi = \psi = 1/\hat{\sigma}$, define $(X', Y') = (\hat{X}(1/\hat{\sigma}), \hat{Y}(1/\hat{\sigma}))$. Dividing both sides of inequalities (8a), (8c), (8e), (8g), (8j) and (8l)

by $\hat{\sigma}$, observe that (X', Y') satisfies (2) with $\hat{\lambda}$ and is therefore in $T_{\text{VRS}}^{\text{R}}$. Because $(\hat{X}, \hat{Y}) = (X'(\hat{\sigma}), Y'(\hat{\sigma}))$, we have $(\hat{X}, \hat{Y}) \in C(I', O') \subseteq \overline{C}(I', O')$.

(ii) Let $\hat{\sigma} = 0$. We now use the following extended notation for any DMU (X, Y):

$$(X,Y) = (X^{V'}, X^{V''}, X^{R'_1}, X^{R'_2}, X^{R''}, Y^{V'}, Y^{V''}, Y^{R'}, Y^{R''}).$$
(22)

In (22), we use the single prime symbol for the measures in the sets I' and O', and the double prime for the measures in I \ I' and O \ O'. For example, the subvector $X^{V'}$ includes inputs $i \in I^V \cap I'$, and $Y^{R''}$ includes outputs $r \in O^R \setminus O'$. The subvector $X^{R'_1}$ includes the inputs from the set $I^R \cap I'$ for which we have a finite bound in (1). The subvector $X^{R'_2}$ includes the inputs from $I^R \cap I'$ for which the bound in (1) is either not specified or is infinite. Because $\hat{\sigma} = 0$, inequalities (8a) and (81) imply that $\hat{Y}^{V'}$ and $\hat{X}^{R'_1}$ are zero vectors. By (8i),

Because $\hat{\sigma} = 0$, inequalities (8a) and (8l) imply that $\hat{Y}^{V'}$ and \hat{X}^{R_1} are zero vectors. By (8i), there exists a $j' \in J$ such that $\hat{\lambda}_{j'} > 0$. Then the inequality (8e) for j = j' implies that $\hat{Y}^{R'}$ is also a zero vector. Using notation as in (22) and replacing the three zero subvectors by **0**, state DMU (\hat{X}, \hat{Y}) as follows:

$$(\hat{X}, \hat{Y}) = (\hat{X}^{V'}, \hat{X}^{V''}, \mathbf{0}, \hat{X}^{R'_2}, \hat{X}^{R''}, \mathbf{0}, \hat{Y}^{V''}, \mathbf{0}, \hat{Y}^{R''}).$$
(23)

Define the sequence of DMUs $\{(X_{(k)}, Y_{(k)})\}, k = 1, 2, ..., as$ follows. Denote $J' = \{j \in J : \hat{\lambda}_j > 0\}$. For each *k*, define

$$\begin{split} X_{(k)}^{V'} &= \sum_{j \in J'} \hat{\lambda}_j X_j^{V'} + k \hat{X}^{V'}, \\ X_{(k)}^{R_1'} &= \overline{\max}_{j \in J'} \{ X_j^{R_1'} \}, \quad X_{(k)}^{R_2'} = \overline{\max}_{j \in J'} \{ X_j^{R_2'} \} + k \hat{X}^{R_2'}, \end{split}$$

where $\overline{\max}$ is the operator of component-wise maximum. For example, each component $X_{(k),i}^{R'_1}$ is defined as the maximum of the corresponding components $X_{i,i}^{R'_1}$, where $j \in J'$.

Let the other components of $(X_{(k)}, Y_{(k)})$ be the same as in (\hat{X}, \hat{Y}) . Note that component $X_{(k)}^{R'_1}$ satisfies (1). It is straightforward to verify that each DMU $(X_{(k)}, Y_{(k)})$ satisfies (2) with the same $\lambda = \hat{\lambda}$ and is therefore in $T_{\text{VRS}}^{\text{R}}$. Using notation (4), where $\varphi = \psi = 1/k$, define $(\hat{X}_{(k)}, \hat{Y}_{(k)}) = (X_{(k)}(1/k), Y_{(k)}(1/k))$, for any k = 1, 2, ... We have:

$$\begin{pmatrix} \hat{X}_{(k)} \\ \hat{Y}_{(k)} \end{pmatrix} = \begin{pmatrix} (1/k) \left(\sum_{j \in J'} \hat{\lambda}_j X_j^{V'} \right) + \hat{X}^{V'}, \, \hat{X}^{V''}, \\ (1/k) \left(\overline{\max}_{j \in J'} \{ X_j^{R'_1} \} \right), \\ (1/k) \left(\overline{\max}_{j \in J'} \{ X_j^{R'_2} \} \right) + \hat{X}^{R'_2}, \, \hat{X}^{R''} \\ (1/k) \hat{Y}^{V'}, \, \hat{Y}^{V''}, \, (1/k) \hat{Y}^{R'}, \, \hat{Y}^{R''} \end{pmatrix} .$$

By definition, $(\hat{X}_{(k)}, \hat{Y}_{(k)}) \in C(I', O')$. DMU (\hat{X}, \hat{Y}) stated in (23) is the limit point of $\{(\hat{X}_{(k)}, \hat{Y}_{(k)})\}$. Therefore, $(\hat{X}, \hat{Y}) \in \overline{C}(I', O')$. Cases (i) and (ii) imply that $W \subseteq \overline{C}(I', O')$. \Box

Lemma 2 If Assumptions 1 and 2 are true, the supremum of program (11) is attained.

Proof of Lemma 2 By Theorem 2, the feasible region Π of program (11) includes all triplets $\langle \lambda, \sigma, \eta \rangle$ that satisfy conditions (8) in which DMU (X, Y) is substituted by $(X_o, Y_o(\eta))$. It suffices to prove that the set Π is closed and bounded. Conditions (8i) and (8n) imply $\lambda \in [0, 1]^n$. As proved in Lemma 1, Assumption 1 implies that there exists a $\bar{\sigma} > 0$ such that $\sigma \in [0, \bar{\sigma}]$. Then condition (8a) stated as $\sum_{j \in J} \lambda_j \sigma Y_{jr}^V \ge \eta Y_{or}^V$, for all $r \in O^V \cap O'$, and Assumption 2 imply that there exists an upper bound M > 0 on η . Therefore, Π is a

bounded set. Furthermore, Π is a closed set because it is defined by nonstrict inequalities and an equality with continuous functions on both sides. Because the objective function η is continuous on Π , its supremum is attained at some $\langle \lambda^*, \sigma^*, \eta^* \rangle \in \Pi$.

Proof of Theorem 3 and Corollary 1 Replace $T_{\text{VRS}}^{\text{R}}$ in program (9) by C(I', O'). For any feasible solution $\langle \varphi, \psi \rangle$ of program (9), we now have the ray of feasible solutions $\langle \alpha \varphi, \alpha \psi \rangle$, $\alpha > 0$, in the modified program. Because $\alpha \psi / \alpha \varphi = \psi / \varphi$, the supremum of ψ / φ is not affected. The ratio ψ / φ is constant along each ray $\langle \alpha \varphi, \alpha \psi \rangle$, $\alpha > 0$, and remains unchanged if we require $\varphi = 1$ and rename ψ as η . Program (9) is now restated as (10).

We have proved that $\theta^* = \eta'$. Let us prove that $\eta' = \eta^*$. Because $C(I', O') \subseteq \overline{C}(I', O')$, we have $\eta' \leq \eta^*$. By Lemma 2, η^* in program (11) is attained. It now suffices to prove that $(X_o, Y_o(\eta^*)) \in C(I', O')$, as stated in Corollary 1.

We first note that $\eta^* \ge 1$. Because $(X_o, Y_o(\eta^*)) \in \overline{C}(I', O')$, by Theorem 2, it satisfies conditions (8) in which we take $(X, Y) = (X_o, Y_o(\eta^*))$, with some $\hat{\lambda}$ and $\hat{\sigma}$. If we assume that $\hat{\sigma} = 0$, conditions (8a) imply that $Y_{or}^V = 0$, $\forall r \in O^V \cap O'$, and (8e) and (8i) imply that $Y_{or}^R = 0$, $\forall r \in O^R \cap O'$, which contradicts Assumption 2. Therefore, $\hat{\sigma} > 0$. Following (4), define DMU $(\tilde{X}, \tilde{Y}) = (X_o(1/\hat{\sigma}), Y_o(\eta^*/\hat{\sigma}))$. Then (\tilde{X}, \tilde{Y}) satisfies (2) with $\hat{\lambda}$ and is in T_{VRS}^R . Because $(X_o, Y_o(\eta^*)) = (\tilde{X}(\hat{\sigma}), \tilde{Y}(\hat{\sigma}))$, we have $(X_o, Y_o(\eta^*)) \in C(I', O')$.

Proof of Theorem 4 By definition (Olesen et al. 2015), technology T_{CRS}^{P} is the intersection of all sets in $\mathbb{R}^{m}_{+} \times \mathbb{R}^{s}_{+}$ that satisfy Axioms 1–4. Similarly, technology T_{CRS} is the intersection of all sets in $\mathbb{R}^{m}_{+} \times \mathbb{R}^{s}_{+}$ that satisfy Axioms 1, 2, 4 and the standard axiom of convexity (Banker et al. 1984). It suffices to prove that, under the assumptions of Theorem 4, Axiom 3 of selective convexity implies the standard axiom of convexity. (The opposite is always true). To be specific, let the single ratio measure be a ratio input. (The case of ratio output is similar and is not considered.) Consider any two DMUs $(\tilde{X}^{V}, \tilde{X}^{R}, \tilde{Y}^{V}) \in T_{CRS}^{P}$ and $(\hat{X}^{V}, \hat{X}^{R}, \hat{Y}^{V}) \in T_{CRS}^{P}$, and any scalar $\gamma \in [0, 1]$. Define

$$(X_{\gamma}^{V}, X_{\gamma}^{R}, Y_{\gamma}^{V}) = \gamma(\tilde{X}^{V}, \tilde{X}^{R}, \tilde{Y}^{V}) + (1 - \gamma)(\hat{X}^{V}, \hat{X}^{R}, \hat{Y}^{V}).$$
(24)

We need to prove that $(X_{\gamma}^V, X_{\gamma}^R, Y_{\gamma}^V) \in T_{CRS}^P$. As assumed, $\tilde{X}^R > 0$ and $\hat{X}^R > 0$. Then $X_{\gamma}^R > 0$, and we can also define the two strictly positive scaling factors

$$\tilde{\alpha} = X_{\gamma}^{R} / \tilde{X}^{R}, \quad \hat{\alpha} = X_{\gamma}^{R} / \hat{X}^{R}.$$
(25)

From (24), we have $\gamma \tilde{X}^R + (1 - \gamma) \hat{X}^R = X_{\gamma}^R$. Dividing both sides of this equality by $X_{\gamma}^R > 0$, noting (25) and rearranging, we have

$$\frac{\gamma}{\tilde{\alpha}} + \frac{(1-\gamma)}{\hat{\alpha}} = 1.$$
(26)

Restate (24) as follows:

$$(X^{V}_{\gamma}, X^{R}_{\gamma}, Y^{V}_{\gamma}) = \frac{\gamma}{\tilde{\alpha}} (\tilde{\alpha} \tilde{X}^{V}, \tilde{\alpha} \tilde{X}^{R}, \tilde{\alpha} \tilde{Y}^{V}) + \frac{(1-\gamma)}{\hat{\alpha}} (\hat{\alpha} \hat{X}^{V}, \hat{\alpha} \hat{X}^{R}, \hat{\alpha} \hat{Y}^{V}).$$
(27)

By Axiom 4, the scaled DMUs $(\tilde{\alpha}\tilde{X}^V, \tilde{\alpha}\tilde{X}^R, \tilde{\alpha}\tilde{Y}^V)$ and $(\hat{\alpha}\hat{X}^V, \hat{\alpha}\hat{X}^R, \hat{\alpha}\hat{Y}^V)$ are in T_{CRS}^P . Note that the scaling is defined in such away that the resulting ratio inputs are now equal. Indeed, from (25), $\tilde{\alpha}\tilde{X}^R = \hat{\alpha}\hat{X}^R = X_{\gamma}^R$. Taking into account (26) and (27), DMU $(X_{\gamma}^V, X_{\gamma}^R, Y_{\gamma}^V)$ is a convex combination of these two scaled DMUs. By Axiom 3, $(X_{\gamma}^V, X_{\gamma}^R, Y_{\gamma}^V) \in T_{CRS}^P$. **Lemma 3** The incorporation of the additional inequality $\eta \ge 1$ in the constraints of programs (13) and (28) (as defined in the proof of Theorem 5) does not affect the value and attainability of their suprema $\hat{\eta}$ and η^* , and does not affect any of their optimal solutions. If Assumption 3 is satisfied, in any of the feasible solutions $\langle \lambda, \sigma, \eta \rangle$ of the resulting programs, we have $\sigma > 0$.

Proof of Lemma 3 The first statement follows from the fact that the solution $\langle \lambda', \sigma', \eta' \rangle$, where λ' is any vector with which DMU $(X_o, Y_o) \in T_{\text{VRS}}^{\text{R}}$ satisfies (2) and $\sigma' = \eta' = 1$, is feasible in (13). Further, assume that in program (13) with the constraint $\eta \ge 1$, we have $\sigma = 0$. Then (13j) and (13o) imply $\lambda = 0$. Then (13b) and (13k) contradict Assumption 3 and the inequality $\eta \ge 1$. Therefore, $\sigma > 0$. The proof for program (28) is similar.

Proof of Theorem 5 Using Theorem 2, we can restate program (11) as follows:

$$\eta^{*} = \max \eta$$
subject to
$$\sum_{j \in J} \lambda_{j} \sigma Y_{jr}^{V} \ge \eta Y_{or}^{V}, \quad \forall r \in O^{V} \cap O',$$

$$\sum_{j \in J} \lambda_{j} \sigma X_{jr}^{V} \ge Y_{or}^{V}, \quad \forall r \in O^{V} \setminus O',$$

$$\sum_{j \in J} \lambda_{j} \sigma X_{ji}^{V} \le X_{oi}^{V}, \quad \forall i \in I^{V} \cap I',$$

$$\sum_{j \in J} \lambda_{j} X_{jr}^{V} = X_{oi}^{V}, \quad \forall i \in I^{V} \setminus I',$$

$$\lambda_{j} (\sigma Y_{jr}^{R} - \eta Y_{or}^{R}) \ge 0, \quad \forall j \in J, \quad \forall r \in O^{R} \cap O',$$

$$\lambda_{j} (\sigma X_{ji}^{R} - X_{oi}^{R}) \le 0, \quad \forall j \in J, \quad \forall r \in O^{R} \setminus O',$$

$$\lambda_{j} (\sigma X_{ji}^{R} - X_{oi}^{R}) \le 0, \quad \forall j \in J, \quad \forall i \in I^{R} \cap I',$$

$$\lambda_{j} (X_{ji}^{R} - X_{oi}^{R}) \le 0, \quad \forall j \in J, \quad \forall i \in I^{R} \setminus I',$$

$$1^{\top} \lambda = 1,$$

$$\eta Y_{or}^{R} \le \sigma \bar{Y}_{r}^{R}, \quad \forall r \in O^{R} \cap O',$$

$$Y_{or}^{R} \le \bar{Y}_{r}^{R}, \quad \forall r \in O^{R} \setminus O',$$

$$X_{oi}^{R} \le \sigma \bar{X}_{i}^{R}, \quad \forall i \in I^{R} \cap I',$$

$$X_{oi}^{R} \le \bar{X}_{i}^{R}, \quad \forall i \in I^{R} \setminus I',$$

$$\lambda \ge \mathbf{0}, \sigma, \eta \ge 0.$$
(28)

By Theorem 3, the optimal value η^* of program (28) is attained. Incorporate the inequality $\eta \ge 1$ as a constraint in programs (13) and (28). Statement (ii) now follows from Lemma 3. Note that, with the constraint $\eta \ge 1$ incorporated (which guarantees $\sigma > 0$), any feasible solution $\langle \lambda^*, \sigma^*, \eta^* \rangle$ of program (28) becomes a feasible solution $\langle \hat{\lambda}, \sigma^*, \eta^* \rangle$ of program (13), where $\hat{\lambda} = \lambda^* \sigma^*$, and vice versa. As noted, η^* is attained. Therefore, $\hat{\eta} = \eta^*$ is also attained in program (13), which proves statement (i). To prove (iii), note that $\langle \hat{\lambda}, \hat{\eta}, \hat{\sigma} \rangle$ is feasible in (13) if and only if $\langle \lambda', \varphi', \psi' \rangle$, where $\lambda' = \hat{\lambda}/\hat{\sigma}, \varphi' = 1/\hat{\sigma}$ and $\psi' = \hat{\eta}/\hat{\sigma}$, is feasible in (9). Because $\psi'/\varphi' = \hat{\eta}$ and, by statement (i), $\hat{\eta} = \theta^*$, solution $\langle \lambda', \varphi', \psi' \rangle$ is optimal in (9) if and only if $\langle \hat{\lambda}, \hat{\eta}, \hat{\sigma} \rangle$ is optimal in (13).

Table 9 DMUs in the example	Measure	Α	В	F	G
	Volume output	1	4	2	3
	Volume input	1	4	2	3
	Ratio output	1	2	2	2
	Ratio input	1	2	2	2

Appendix B: Comparison of cone $\bar{C}(I, O)$ and technology T_{CBS}^{P}

The following example clarifies the difference between the full closed cone extension $\bar{C}(I, O)$ of technology $T_{\text{VRS}}^{\text{R}}$ and technology $T_{\text{CRS}}^{\text{P}}$ of Olesen et al. (2015). It shows that (i) $\overline{C}(I, O)$ does not satisfy Axiom 3 of selective convexity and (ii) $\overline{C}(I, O) \neq T_{\text{CRS}}^{\text{P}}$. For simplicity, we assume that there are no upper bounds (1) on the ratio input and output.

Let DMUs A and B in Table 9 be observed. DMU F is not observed and is obtained by

scaling DMU A with $\alpha = 2$. All three DMUs A, B and F belong to $\overline{C}(I, O)$ and to T_{CRS}^{P} . By definition, technology T_{CRS}^{P} satisfies Axiom 3 of selective convexity. Because both ratio input and output of DMUs B and F are equal, by Axiom 3, DMU G, constructed as the average of DMUs B and F, belongs to technology T_{CRS}^P . Let us show that $G \notin \overline{C}(I, O)$ and, therefore, the closed cone $\bar{C}(I, O)$ does not satisfy Axiom 3 of selective convexity.

We need to prove that conditions (8) in which I' = I and O' = O and DMU (X, Y) is G cannot be satisfied. Changing notation from λ_1 and λ_2 to λ_A and λ_B , we have:

$$1\sigma\lambda_A + 4\sigma\lambda_B \ge 3,\tag{29a}$$

$$1\sigma\lambda_A + 4\sigma\lambda_B \le 3,\tag{29b}$$

$$\lambda_A(1\sigma - 2) \ge 0, \tag{29c}$$

$$\lambda_A(1\sigma - 2) \le 0,\tag{29d}$$

$$\lambda_B(2\sigma - 2) \ge 0, \tag{29e}$$

$$\lambda_B(2\sigma - 2) \le 0,\tag{29f}$$

$$\lambda_A + \lambda_B = 1, \tag{29g}$$

$$\lambda_A, \lambda_B, \sigma \ge 0. \tag{29h}$$

Let $\lambda_A = 1$ and $\lambda_B = 0$. Then (29a) and (29b) imply $\sigma = 3$. However, (29c) and (29d) imply $\sigma = 2$. Therefore, this case is impossible. Similarly, if $\lambda_A = 0$ and $\lambda_B = 1$, then (29a) and (29b) imply $\sigma = 3/4$, while (29e) and (29f) imply $\sigma = 1$. Therefore, this case is also impossible. Finally, assume that $\lambda_A > 0$ and $\lambda_B > 0$. Then (29c) and (29d) imply $\sigma = 2$. This is inconsistent with conditions (29e) and (29f) which imply $\sigma = 1$.

We have shown that DMU $G \notin C(I, O)$. Therefore, the cone C(I, O) does not satisfy Axiom 3 of selective convexity. In contrast, technology T_{CRS}^{P} satisfies Axiom 3 by definition (Olesen et al. 2015). In particular, DMU G satisfies (12) with $\lambda_A = \lambda_B = 1/2$, $\sigma_A = 2$ and $\sigma_B = 1$. Therefore, $G \in T_{CRS}^P$. This also implies that $\bar{C}(I, O) \neq T_{CRS}^P$.

Appendix C: The data set used in the application is Sect. 8

Table 10 shows the inputs and outputs of all schools used in the application in Sect. 8. The latter section includes description of the inputs and outputs. Summary statistics for these data were given in Table 5.

Table 10 The data set used in the application	School	Teachers	Expenditure	NFSM	Pupils	HE
••	1	70.2	3203.4	66.1	1013	0.72
	2	93.7	3425.8	55.7	1225	0.75
	3	124.9	5239.7	68.7	1830	0.82
	4	85.3	4472.2	57.2	1434	0.91
	5	40.8	1778.7	81.5	1010	0.85
	6	90.5	5638.9	76.9	1242	0.72
	7	71.8	3184	64.7	1184	0.69
	8	101	4297.3	71.2	1536	0.87
	9	57	2120.8	86.3	1133	0.63
	10	72.6	4023.1	74.1	1477	0.84
	11	64.3	3212	57.5	1004	0.71
	12	57.6	3260.3	84.8	888	0.55
	13	71.5	2743.6	88.1	1208	0.74
	14	76	2425.8	90	1255	0.68
	15	34.4	1253.1	76.1	503	0.6
	16	73.8	2655.5	79.6	1059	0.55
	17	95.5	2949	90.2	1429	0.62
	18	83.6	4454.1	76.9	1409	0.5
	19	60.5	2728.2	81.3	963	0.83
	20	53.8	2151.4	71.1	765	0.73
	21	47.3	1620	84	746	0.7
	22	64.6	2773.8	96.1	1070	0.76
	23	71.6	2425.2	96.8	1199	0.73
	24	80.3	3185.7	92.9	1399	0.83
	25	63.5	2543.9	95.4	1084	0.83
	26	49	2011.8	81.8	771	0.76
	27	29.9	1216.7	80	490	0.45
	28	63.9	1960.6	90.7	1027	0.6
	29	63.7	2406.3	89.5	1076	0.55
	30	84	3060.3	94.3	1505	0.75
	31	38.6	1612.4	75.8	628	0.6
	32	45.7	2283	80	769	0.74
	33	22.2	1184	64.1	396	0.5
	34	48.6	1977.5	87.2	847	0.61
	35	45.1	1724.2	61.2	678	0.87
	36	62.4	2019.6	86.1	968	0.48
	37	55.1	2154.6	84.3	814	0.52
	38	116.4	5275.7	68.9	1933	0.69
	39	48	1521.6	86.7	766	0.8

References

- Afriat, S. N. (1972). Efficiency estimation of production functions. *International Economic Review*, 13(3), 568–598.
- Banker, R. D. (1984). Estimating most productive scale size using data envelopment analysis. *European Journal of Operational Research*, 17(1), 35–44.
- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9), 1078–1092.
- Banker, R. D., & Morey, R. C. (1986). Efficiency analysis for exogenously fixed inputs and outputs. Operations Research, 34(4), 513–521.
- Banker, R. D., & Thrall, R. M. (1992). Estimation of returns to scale using data envelopment analysis. European Journal of Operational Research, 62(1), 74–84.
- Bixby, R. E. (2002). Solving real-world linear programs: A decade or more of progress. Operations Research, 50(1), 3–15.
- Chambers, R. G., & Färe, R. (2008). A "calculus" for data envelopment analysis. Journal of Productivity Analysis, 30(3), 169–175.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444.
- Cooper, W. W., Pastor, J. T., Borras, F., Aparicio, J., & Pastor, D. (2011). BAM: A bounded adjusted measure of efficiency for use with bounded additive models. *Journal of Productivity Analysis*, 35(2), 85–94.
- Cooper, W. W., Seiford, L. M., & Tone, K. (2007). Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software (2nd edn.). Springer.
- Deprins, D., Simar, L., & Tulkens, H. (1984). Measuring labor-efficiency in post offices. In M. Marchand, P. Pestieau, & H. Tulkens (Eds.), *The performance of public enterprises: Concepts and measurements* (pp. 243–267). North-Holland.
- Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., & Shale, E. A. (2001). Pitfalls and protocols in DEA. European Journal of Operational Research, 132(2), 245–259.
- Elble, J. M., & Sahinidis, N. V. (2012). Scaling linear optimization problems prior to application of the simplex method. *Computational Optimization and Applications*, 52(2), 345–371.
- Emrouznejad, A., & Amin, G. R. (2009). DEA models for ratio data: Convexity consideration. Applied Mathematical Modelling, 33(1), 486–498.
- Färe, R., Grosskopf, S., & Logan, J. (1983a). The relative efficiency of Illinois electric utilities. *Resources and Energy*, 5(4), 349–367.
- Färe, R., Grosskopf, S., & Lovell, C. A. K. (1983b). The structure of technical efficiency. Scandinavian Journal of Economics, 85(2), 181–190.
- Färe, R., Grosskopf, S., & Lovell, C. A. K. (1985). The measurement of efficiency of production. Kluwer Academic Publishers.
- Färe, R., & Lovell, C. A. K. (1978). Measuring the technical efficiency of production. *Journal of Economic Theory*, 19(1), 150–162.
- Førsund, F. R., & Hjalmarsson, L. (2004). Calculating scale elasticity in DEA models. *Journal of the Opera*tional Research Society, 55(10), 1023–1038.
- Frisch, R. (1965). Theory of production. D. Reidel Publishing Company.
- Golany, B., & Roll, Y. (1993). Some extensions of techniques to handle non-discretionary factors in data envelopment analysis. *Journal of Productivity Analysis*, 4(4), 419–432.
- Hadjicostas, P., & Soteriou, A. C. (2006). One-sided elasticities and technical efficiency in multi-output production: A theoretical framework. *European Journal of Operational Research*, 168(2), 425–449.
- Mostafaee, A., & Soleimani-damaneh, M. (2020). Global sub-increasing and global sub-decreasing returns to scale in free disposal hull technologies: Definition, characterization and calculation. *European Journal* of Operational Research, 280(1), 230–241.
- Olesen, O. B., Petersen, N. C., & Podinovski, V. V. (2015). Efficiency analysis with ratio measures. *European Journal of Operational Research*, 245(2), 446–462.
- Olesen, O. B., Petersen, N. C., & Podinovski, V. V. (2017). Efficiency measures and computational approaches for data envelopment analysis models with ratio inputs and outputs. *European Journal of Operational Research*, 261(2), 640–655.
- Olesen, O. B., Petersen, N. C., & Podinovski, V. V. (2022). The structure of production technologies with ratio inputs and outputs. *Journal of Productivity Analysis*, 57(3), 255–267. https://doi.org/10.1007/s11123-022-00631-6.
- Pastor, J. T., Aparicio, J., Alcaraz, J., Vidal, F., & Pastor, D. (2015). An enhanced BAM for unbounded or partially bounded CRS additive models. *Omega*, 56, 16–24.

- 423
- Pastor, J. T., Aparicio, J., Monge, J. F., & Pastor, D. (2013). Modeling CRS bounded additive DEA models and characterizing their Pareto-efficient points. *Journal of Productivity Analysis*, 40(3), 285–292.
- Podinovski, V. V. (2004a). Efficiency and global scale characteristics on the "no free lunch" assumption only. *Journal of Productivity Analysis*, 22(3), 227–257.
- Podinovski, V. V. (2004b). Local and global returns to scale in performance measurement. *Journal of the Operational Research Society*, 55(2), 170–178.
- Podinovski, V. V. (2005). Selective convexity in DEA models. European Journal of Operational Research, 161(2), 552–563.
- Podinovski, V. V. (2017). Returns to scale in convex production technologies. *European Journal of Operational Research*, 258(3), 970–982.
- Podinovski, V. V., Chambers, R. G., Atici, K. B., & Deineko, I. D. (2016). Marginal values and returns to scale for nonparametric production frontiers. *Operations Research*, 64(1), 236–250.
- Ruggiero, J. (1996). On the measurement of technical efficiency in the public sector. European Journal of Operational Research, 90(3), 553–565.
- Sahoo, B. K., & Tone, K. (2015). Scale elasticity in non-parametric DEA approach. In J. Zhu (Ed.), Data envelopment analysis: A handbook of models and methods (pp. 269–290). Springer.
- Shephard, R. W. (1974). Indirect production functions. Mathematical Systems in Economics. Anton Hain.
- Zelenyuk, V. (2013). A scale elasticity measure for directional distance function and its dual: Theory and DEA estimation. *European Journal of Operational Research*, 228(3), 592–600.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.