

Efficiency measures and computational approaches for data envelopment analysis models with ratio inputs and outputs

Ole Bent Olesen^{a,*}, Niels Christian Petersen^a, Victor V. Podinovski^b

^a*Department of Business Economics, The University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark*

^b*School of Business and Economics, Loughborough University, Loughborough LE11 3TU, UK*

Abstract

In a recent paper to this journal, the authors developed a methodology that allows the incorporation of ratio inputs and outputs in the variable and constant returns-to-scale DEA models. Practical evaluation of efficiency of decision making units (DMUs) in such models generally goes beyond the application of standard linear programming techniques. In this paper we discuss how the DEA models with ratio measures can be solved. We also introduce a new type of potential ratio (PR) inefficiency. It characterizes DMUs that are strongly efficient in the model of technology with ratio measures but become inefficient if the volume data used to calculate ratio measures become available. Potential ratio inefficiency can be tested by the programming approaches developed in this paper.

Keywords: data envelopment analysis, ratio measures, efficiency

1. Introduction

The two conventional models of data envelopment analysis (DEA) are stated under the assumptions of variable and constant (VRS and CRS) returns-to-scale (Charnes et al. 1978, Banker et al. 1984). In both models the underlying production technology is assumed to satisfy several known axioms (Banker et al. 1984).

In a recent paper to this journal (Olesen et al. 2015) we showed that the use of ratio inputs and outputs (such as percentages or various rates) in the VRS and CRS models generally violates the stated production assumptions. In particular, if ratio measures are used, the underlying technology generally becomes nonconvex. Therefore, the standard DEA models are generally not suitable if at least one input or output is a ratio.

In the same paper, we introduced new Ratio-VRS (R-VRS) and Ratio-CRS (R-CRS) production technologies that can be regarded as extensions to the standard DEA models that allow the incorporation of ratio inputs and outputs. The R-VRS model is also an extension of the approach developed by Ruggiero (1996) to the treatment of exogenous inputs that are used to control for environmental factors. Applications of the latter approach have been reported, for example, in Blackburn et al. (2014), Haelermans and Ruggiero (2013), and Johnson and Ruggiero (2014). The R-CRS model provides a further extension assuming the scalability of volume inputs and outputs, like in the standards CRS model, while excluding ratio measures from this assumption.

*Corresponding author

Email addresses: ole@sam.sdu.dk (Ole Bent Olesen), ncp@sam.sdu.dk (Niels Christian Petersen), v.podinovski@lboro.ac.uk (Victor V. Podinovski)

In the above paper (Olesen et al. 2015), we also outlined a computational procedure that could be used to assess the efficiency of decision making units (DMUs) in the R-VRS and R-CRS technologies.¹ Depending on the measures to be improved (e.g., only volume or ratio inputs, or both) and also on the assumption of returns to scale (VRS or CRS), the solution procedure requires the use of linear programming, mixed integer linear programming or nonlinear optimization techniques.

Since the publication of the above paper, it has been brought to our attention that it would be useful to clarify the suggested solution procedures. This paper addresses this gap and provides a more detailed discussion important for practical application of the R-VRS and R-CRS DEA models.

In particular, we discuss solution approaches for the assessment of radial and nonradial efficiency in the R-VRS technology, and also in the practically important case of R-CRS technology with fixed ratio inputs and outputs. As in the case of standard VRS and CRS models, all such efficiency measures usually result in a weakly efficient projection of the DMU under the assessment on the boundary of technology. We show how the second optimization stage of Ali and Seiford (1993) maximizing the sum of input and output slacks, and resulting in a strongly efficient target DMU, can be implemented in technologies with ratio inputs and outputs.

In this paper we also identify a new type of inefficiency that has no direct analogue in the standard VRS and CRS technologies, referred to as potential ratio (PR) inefficiency. A PR-inefficient DMU is strongly efficient in technology with ratio measures but becomes inefficient if the volume data (numerator and denominator) used in the calculation of ratio inputs and outputs become available. We illustrate this new type of inefficiency by simple examples and develop computational approaches for its testing.

We complete our development by a numerical example that illustrates the computational approaches developed in this paper, and demonstrates the usefulness of the new notion of PR efficiency in interpreting optimal solutions to DEA models with ratio inputs and outputs.

2. Preliminaries

Let T be a production technology with m inputs and s outputs. Denote I^V and O^V the subsets of volume (V) inputs and outputs, respectively. Similarly, let I^R and O^R be the subsets of ratio (R) inputs and outputs.

Following Olesen et al. (2015), each DMU (X, Y) in technology T is stated in the form

$$(X, Y) = (X^V, X^R, Y^V, Y^R),$$

where $X = (X^V, X^R) \in \mathbb{R}_+^m$ and $Y = (Y^V, Y^R) \in \mathbb{R}_+^s$ are the vectors of volume and ratio inputs and outputs, respectively. Denote observed DMUs (X_j, Y_j) , where $j \in J = \{1, \dots, n\}$.

Olesen et al. (2015) show that, if some inputs or outputs are ratio measures, the standard axioms of Banker et al. (1984), on which the VRS technology is based, need modifying. In particular, the axiom of convexity is replaced by the axiom of *selective convexity* (Podinovski 2005). This axiom assumes that convex combinations of DMUs are possible provided these DMUs have *identical* vectors of ratio inputs and outputs.

If the vectors of ratio inputs and outputs of the combined DMUs are not identical, their aggregation is still possible but requires a different treatment of volume and ratio measures.

¹Any feasible combination of inputs and outputs within the given technology is referred to as a DMU. Where specifically indicated, we distinguish between observed and unobserved DMUs.

While the volume inputs and outputs are aggregated exactly as in the standard convex combination, each ratio input is taken at the highest level across all DMUs forming such combination, and each ratio output is taken at the lowest level.²

The reason for this special treatment of ratio inputs and outputs is that such measures cannot be aggregated with the same weights as the volume measures. For the correct aggregation we need to know the numerator and denominator of each ratio input and output, e.g., as assumed in Emrouznejad and Amin (2009). Olesen et al. (2015) do not assume that such information is available and, in aggregating ratio inputs and output, allow for the most conservative worst-case scenario that may arise if the volume numerators and denominators become known.

The described treatment means that, in models with ratio measures, the standard notion of a convex combination of DMUs is replaced by the notion of *ratio-convex combination* of DMUs introduced next.

Consider any finite number of DMUs $(\tilde{X}_k, \tilde{Y}_k) \in T$, $k \in \mathcal{K} = \{1, \dots, K\}$, and any vector $\lambda \in \mathbb{R}^K$ such that $\lambda_k > 0$, $\forall k$, and $\sum_{k \in \mathcal{K}} \lambda_k = 1$. Define DMU $(\hat{X}, \hat{Y}) = (\hat{X}^V, \hat{X}^R, \hat{Y}^V, \hat{Y}^R)$ as follows:

$$\hat{Y}^V = \sum_{k \in \mathcal{K}} \lambda_k \tilde{Y}_k^V, \quad (1.1)$$

$$\hat{X}^V = \sum_{k \in \mathcal{K}} \lambda_k \tilde{X}_k^V, \quad (1.2)$$

$$\hat{Y}_r^R = \min_{k \in \mathcal{K}} \left\{ \tilde{Y}_{kr}^R \right\}, \quad \forall r \in \mathcal{O}^R, \quad (1.3)$$

$$\hat{X}_i^R = \max_{k \in \mathcal{K}} \left\{ \tilde{X}_{ki}^R \right\}, \quad \forall i \in \mathcal{I}^R. \quad (1.4)$$

Definition 1. DMU (\hat{X}, \hat{Y}) is referred to as a ratio-convex (R-convex) combination of DMUs $(\tilde{X}_k, \tilde{Y}_k)$, $k \in \mathcal{K}$.³

Note that, in the above definition, we assume that $\lambda_k > 0$, for all k . Of course, we can assume a more general condition $\lambda_k \geq 0$, for all k , in which case the minimum in equality (1.3) and the maximum in (1.4) should be taken over all $k \in \mathcal{K}$ such that $\lambda_k > 0$.

Based on the new set of production axioms that allows R-convex combinations of DMUs, Olesen et al. (2015) develop the R-VRS technology T_{VRS}^R with volume and ratio inputs and outputs. This technology coincides with the set of all DMUs $(X, Y) \in \mathbb{R}_+^m \times \mathbb{R}_+^s$ for which

²This type of aggregation of volume and ratio inputs and outputs is valid in any technology T that satisfies the axioms of selective convexity and free disposability of ratio inputs and outputs (Podinovski 2005, Axiom A3' and Lemma 2).

³DMUs $(\tilde{X}_k, \tilde{Y}_k)$ may be observed or unobserved. A R-convex combination of DMUs is introduced in Axiom A3' of Podinovski (2005), without using this specific term.

there exists a vector $\lambda \in \mathbb{R}^n$ such that⁴

$$\begin{aligned}
\sum_{j \in J} \lambda_j Y_j^V &\geq Y^V, \\
\sum_{j \in J} \lambda_j X_j^V &\leq X^V, \\
\lambda_j (Y_j^R - Y^R) &\geq \mathbf{0}, \quad \forall j \in J, \\
\lambda_j (X_j^R - X^R) &\leq \mathbf{0}, \quad \forall j \in J, \\
\mathbf{1}^\top \lambda &= 1, \\
\lambda &\geq \mathbf{0}.
\end{aligned} \tag{2}$$

In situations where there are natural upper bounds on some ratio inputs and outputs (e.g., 1 or 100%), these are also incorporated in the definition of technology $T_{\text{VRS}}^{\text{R}}$. However, as proved in Proposition 7 in Olesen et al. (2015), any such bounds would be automatically satisfied in any practical assessment of efficiency discussed below and are therefore omitted in our development.

Furthermore, Olesen et al. (2015) develop the R-CRS technology $T_{\text{CRS}}^{\text{R}}$ assuming that the volume inputs and outputs can be changed proportionally with an arbitrary scaling factor $\alpha \geq 0$, as in the standard CRS technology. At the same time, the ratio inputs and outputs may respond to this scaling in different ways. Olesen et al. (2015) identify different types of ratio measures: fixed, proportional, downward-proportional and upward-proportional. In a formal statement of technology $T_{\text{CRS}}^{\text{R}}$, each of these types of ratio measure requires a different treatment and is expressed by different mathematical conditions. A formal statement of technology $T_{\text{CRS}}^{\text{R}}$ is given by Olesen et al. (2015) and is not reproduced in this paper.

3. Solving R-VRS models

3.1. Assessing the input radial efficiency

Consider assessing the input radial efficiency of some DMU $(X_o, Y_o) \in T_{\text{VRS}}^{\text{R}}$.⁵ To simplify the exposition and at the same time provide an illustration to the general case, suppose we wish to measure the efficiency with respect to all volume and ratio inputs. In practice, however, we may wish to measure the input radial efficiency with respect to a subset of inputs, e.g., volume inputs only, while keeping the environmental ratio inputs fixed. The latter case is well-known in the literature. It corresponds to the model proposed in Ruggiero (1996) and has a clear economic interpretation.⁶

Using conditions (2), the input radial efficiency of DMU $(X_o, Y_o) \in T_{\text{VRS}}^{\text{R}}$, also stated as

⁴The conditions for ratio inputs and outputs in (2) can equivalently be restated as follows: if $\lambda_j > 0$ then $Y_j^R \geq Y^R$ and $X_j^R \leq X^R$, for all $j \in J$.

⁵DMU (X_o, Y_o) may be observed or unobserved. Throughout the paper we assume that $X_o \neq \mathbf{0}$ and $Y_o \neq \mathbf{0}$ without further mention.

⁶As highlighted in Olesen et al. (2015, p. 455) in some situations “we may also be interested in possible improvements to ratio measures. In such cases θ may not have the conventional meaning of input radial efficiency, but it is still a measure of input improvement”. The two examples in footnote 19 in Olesen et al. (2015) outline such situations.

$(X_o^V, X_o^R, Y_o^V, Y_o^R)$, is the optimal value θ^* of the following program:

$$\theta^* = \min \quad \theta \tag{3.1}$$

$$\text{s.t.} \quad \sum_{j \in J} \lambda_j Y_j^V \geq Y_o^V, \tag{3.2}$$

$$\sum_{j \in J} \lambda_j X_j^V \leq \theta X_o^V, \tag{3.3}$$

$$\lambda_j (Y_j^R - Y_o^R) \geq \mathbf{0}, \quad \forall j \in J, \tag{3.4}$$

$$\lambda_j (X_j^R - \theta X_o^R) \leq \mathbf{0}, \quad \forall j \in J, \tag{3.5}$$

$$\mathbf{1}^\top \lambda = 1, \tag{3.6}$$

$$\lambda \geq \mathbf{0}, \theta \text{ sign free.} \tag{3.7}$$

In the above program the vector inequalities (3.5) are nonlinear. However, these inequalities are easy to linearize. Indeed, we first restate (3.5) in the equivalent form:

$$\text{either } \lambda_j = 0 \text{ or } X_j^R - \theta X_o^R \leq \mathbf{0}, \quad \forall j \in J, \tag{4}$$

and introduce a variable vector $\delta \in \{0, 1\}^n$ with binary components. The “either-or” condition (4) is linearized by restating it as the two inequalities:

$$\lambda_j \leq \delta_j, \quad \forall j \in J, \tag{5.1}$$

$$X_j^R - \theta X_o^R \leq L_1(1 - \delta_j), \quad \forall j \in J, \tag{5.2}$$

where L_1 is a constant vector of dimension $|I^R|$ with sufficiently large positive components.⁷ This transforms model (3) to a mixed integer linear program.

A practical way to define a suitable vector L_1 (and similar vectors for the other programs stated below) is discussed in Appendix B.

3.2. Assessing the output radial efficiency

The output radial efficiency of DMU (X_o, Y_o) is the inverse of the optimal value η^* in the following program:

$$\eta^* = \max \quad \eta \tag{6.1}$$

$$\text{s.t.} \quad \sum_{j \in J} \lambda_j Y_j^V \geq \eta Y_o^V, \tag{6.2}$$

$$\sum_{j \in J} \lambda_j X_j^V \leq X_o^V, \tag{6.3}$$

$$\lambda_j (Y_j^R - \eta Y_o^R) \geq \mathbf{0}, \quad \forall j \in J, \tag{6.4}$$

$$\lambda_j (X_j^R - X_o^R) \leq \mathbf{0}, \quad \forall j \in J, \tag{6.5}$$

$$\mathbf{1}^\top \lambda = 1, \tag{6.6}$$

$$\lambda \geq \mathbf{0}, \eta \text{ sign free.} \tag{6.7}$$

⁷Theoretically, components of vector L_1 may be very large. However, using an arbitrary *excessively* large vector L_1 may create well-known computational problems (see Appendix A for details). It would therefore be advisable to choose vector L_1 , and the other “large” vectors in this paper, close to the reasonably conservative theoretical estimates obtained in Appendix B.

The nonlinear constraints (6.4) of this program can be linearized using the same technique as in the case of input minimization. Indeed, restate the inequality (6.4) in the equivalent form:

$$\text{either } \lambda_j = 0 \text{ or } \eta Y_o^R - Y_j^R \leq \mathbf{0}, \quad \forall j \in J.$$

We now replace the above "either-or" condition by the linear inequalities

$$\lambda_j \leq \delta_j, \quad \forall j \in J, \quad (7.1)$$

$$\eta Y_o^R - Y_j^R \leq L_2(1 - \delta_j), \quad \forall j \in J, \quad (7.2)$$

where $\delta \in \{0, 1\}^n$ is a variable vector, and L_2 is a constant vector of dimension $|O^R|$ with sufficiently large components. This transforms model (6) to a mixed integer linear program.

3.3. Nonradial measures

The described approach extends to a more general case in which, for example, the inputs are improved in a nonradial direction $g_X = (g_X^V, g_X^R) \geq \mathbf{0}$, $g_X \neq \mathbf{0}$ (Chambers et al. 1996, 1998). In this case the improved DMU is described parametrically as

$$(X(\beta), Y_o) = (X_o^V - \beta g_X^V, X_o^R - \beta g_X^R, Y_o^V, Y_o^R), \quad (8)$$

where $\beta \geq 0$ is a scalar variable, and program (3) is replaced by

$$\beta^* = \max \beta \quad (9.1)$$

$$\text{s.t. } \sum_{j \in J} \lambda_j Y_j^V \geq Y_o^V, \quad (9.2)$$

$$\sum_{j \in J} \lambda_j X_j^V \leq X_o^V - \beta g_X^V, \quad (9.3)$$

$$\lambda_j (Y_j^R - Y_o^R) \geq \mathbf{0}, \quad \forall j \in J, \quad (9.4)$$

$$\lambda_j (X_j^R - (X_o^R - \beta g_X^R)) \leq \mathbf{0}, \quad \forall j \in J, \quad (9.5)$$

$$\mathbf{1}^\top \lambda = 1, \quad (9.6)$$

$$\lambda \geq \mathbf{0}, \beta \geq 0. \quad (9.7)$$

Similar to the above treatment, the nonlinear constraints (9.5) can be restated in the equivalent form:

$$\text{either } \lambda_j = 0 \text{ or } X_j^R - (X_o^R - \beta g_X^R) \leq \mathbf{0}, \quad \forall j \in J.$$

These conditions are further linearized as

$$\lambda_j \leq \delta_j, \quad \forall j \in J, \quad (10.1)$$

$$X_j^R - (X_o^R - \beta g_X^R) \leq L_3(1 - \delta_j), \quad \forall j \in J, \quad (10.2)$$

where $\delta \in \{0, 1\}^n$ is a variable vector, and L_3 is a sufficiently large constant vector of dimension $|I^R|$.

4. Solving R-CRS models

Olesen et al. (2015) show that assessing the input or output radial efficiency of DMU (X_o, Y_o) in technology $T_{\text{CRS}}^{\text{R}}$ generally requires solving a nonlinear program, and suggest a simple line search algorithm for this purpose. Below we comment on a special case of technology $T_{\text{CRS}}^{\text{R}}$ which is important from a practical point of view and for which a simpler solution method can be identified.

Suppose that all ratio inputs and outputs are of the *fixed type*. These measures can be assumed to remain constant when the volume measures are scaled with some factor α . Examples of this type include contextual measures such as non-discretionary socio-economic factors, and also quality factors that should not change with the volume of production.

Following Olesen et al. (2015), let $T_{\text{CRS}}^{\text{F}}$ denote the R-CRS technology in which all ratio inputs and outputs are of the fixed type.⁸ According to Proposition 4 in Olesen et al. (2015), the mathematical statement of this technology is obtained from the statement (2) of technology $T_{\text{VRS}}^{\text{R}}$ by removing the normalizing condition $\mathbf{1}^\top \lambda = 1$.⁹

To be specific, consider assessing the input radial efficiency DMU $(X_o, Y_o) \in T_{\text{CRS}}^{\text{F}}$. For this we solve program (3) from which the normalizing equality (3.6) is removed.

As in the case of model (3), we restate the nonlinear constraints (3.5) by the “either-or” condition (4). Note, however, that the latter cannot be linearized by conditions (5) because variables λ_j are no longer bounded above by 1, and the inequality (5.1) is generally unsuitable.

We overcome this difficulty by an additional step. First note that we can limit the feasible values of θ by the condition $\theta \leq 1$. Then for all $j \in J$, from (3.3) we have $\lambda_j \leq \Lambda_j$, where

$$\Lambda_j = \min\{X_{oi}^V/X_{ji}^V \mid i \in I^V : X_{ji}^V > 0\}, \quad \forall j \in J.$$

Define $\Lambda = \max_{j \in J} \Lambda_j$. Note that no λ_j in a feasible solution can exceed Λ . We now replace inequality (5.1) by the following condition:

$$\lambda_j \leq \Lambda \delta_j, \quad \forall j \in J. \quad (11)$$

The nonlinear inequality (3.5) is now replaced by the two linear inequalities (5.2) and (11), using the same constant vector L_1 as above. This transforms the original nonlinear program to a mixed integer linear form.

The cases of output radial efficiency and nonradial efficiency improvements with a chosen directional vector are treated in a similar way.

5. The second optimization stage

It is well known that assessing the input or output radial efficiency of DMU (X_o, Y_o) in the standard VRS or CRS technology projects it on the boundary of the technology, but the projection may not be (strongly) efficient. To test if a DMU is strongly or only weakly efficient, the second optimization stage of Ali and Seiford (1993) is usually performed.¹⁰

⁸Olesen et al. (2015) note that technology $T_{\text{CRS}}^{\text{F}}$ may be seen as the partial cone extension of technology $T_{\text{VRS}}^{\text{R}}$.

⁹As stated in Proposition 4 in Olesen et al. (2015), conditions (2) without the normalizing condition $\mathbf{1}^\top \lambda = 1$ describe all DMUs $(X, Y) \in T_{\text{CRS}}^{\text{F}}$ whose vectors of volume outputs are nonzero, i.e., $Y^V \neq \mathbf{0}$. This assumption should not be restrictive in most practical applications. The general statement of technology $T_{\text{CRS}}^{\text{F}}$, without the assumption $Y^V \neq \mathbf{0}$, is more complex and is given by Theorem 3 in Olesen et al. (2015).

¹⁰Often the two optimization stages are combined in one. In this approach the input and output slacks are incorporated in the objective function with a very small positive multiplier ε .

A similar second optimization stage is applicable to the R-VRS and R-CRS models of Olesen et al. (2015). To be specific, below we first consider the R-VRS model, and comment on the R-CRS model at the end of this section.

We start with two general definitions applicable to any production technology T .

Definition 2. DMU $(X, Y) \in T$ is (strongly) efficient if there does not exist a DMU $(X', Y') \in T$, such that $X' \leq X$, $Y' \geq Y$, and $(X', Y') \neq (X, Y)$.

Restating Definition 2, DMU (X, Y) is strongly efficient if

$$\{(X', Y') \in T \mid X' \leq X, Y' \geq Y\} \cap T = \{(X, Y)\}.$$

Definition 3. A DMU $(X, Y) \in T$ is weakly efficient if there does not exist a DMU $(X', Y') \in T$, such that $X' < X$ and $Y' > Y$.¹¹

Obviously, a strongly efficient DMU is weakly efficient but the converse is not true.

Let (\hat{X}, \hat{Y}) be the projection of DMU (X_o, Y_o) on the boundary of the R-VRS technology $T_{\text{VRS}}^{\text{R}}$. Namely, in the input oriented model (3), we have $(\hat{X}, \hat{Y}) = (\theta^* X_o, Y_o)$. In the output oriented model (6), $(\hat{X}, \hat{Y}) = (X_o, \eta^* Y_o)$. For the nonradial model (9), DMU (\hat{X}, \hat{Y}) is defined by formula (8), where we take $\beta = \beta^*$.¹²

It is clear that the projection (\hat{X}, \hat{Y}) is weakly efficient in $T_{\text{VRS}}^{\text{R}}$.¹³ To test if the projected DMU (\hat{X}, \hat{Y}) is strongly efficient, we apply the second optimization stage of Ali and Seiford (1993) and solve the following program that maximizes the sum of input and output slacks, i.e., components of vectors $\xi \in \mathbb{R}_+^m$ and $\zeta \in \mathbb{R}_+^s$.¹⁴

$$\begin{aligned} \Delta^* = \max \quad & \sum_{i=1}^m \xi_i + \sum_{r=1}^s \zeta_r \\ \text{s.t.} \quad & (\hat{X} - \xi, \hat{Y} + \zeta) \in T_{\text{VRS}}^{\text{R}}, \\ & \xi, \zeta \geq \mathbf{0}. \end{aligned} \tag{12}$$

Let $(\hat{\xi}, \hat{\zeta})$ be an optimal solution to program (12). Define DMU

$$(X^*, Y^*) = (\hat{X} - \hat{\xi}, \hat{Y} + \hat{\zeta}). \tag{13}$$

DMU (X^*, Y^*) is often referred to as an *efficient target* of DMU (X_o, Y_o) .¹⁵

Proposition 1. DMU (X^*, Y^*) is strongly efficient in technology $T_{\text{VRS}}^{\text{R}}$.

Proposition 2. DMU (X_o, Y_o) is strongly efficient in technology $T_{\text{VRS}}^{\text{R}}$ if and only if the two conditions are satisfied:

- (i) The first-stage projection (\hat{X}, \hat{Y}) coincides with DMU (X_o, Y_o) , i.e., $\theta^* = 1$ in the input-oriented model (3), or $\eta^* = 1$ in the output-oriented model (6), or $\beta^* = 0$ in the nonradial model (9).
- (ii) The maximum sum of slacks Δ^* in program (12) is equal to zero.

¹¹The strict inequality between vectors means that this inequality is true for each component. Thus, $X' < X$ means $X'_i < X_i, \forall i = 1, \dots, m$.

¹²The DMU (\hat{X}, \hat{Y}) is a boundary point of technology $T_{\text{VRS}}^{\text{R}}$ because any further improvement of DMU (X_o, Y_o) in the chosen direction (e.g., any radial improvement of the input vector X_o with $\theta < \theta^*$ in the input oriented model (3)) leads outside technology $T_{\text{VRS}}^{\text{R}}$.

¹³Otherwise, by Definition 3, there exists a DMU $(X', Y') \in T_{\text{VRS}}^{\text{R}}$, such that $X' < X^*$ and $Y' > Y^*$, which contradicts the definition of projection and optimality of the improvement factor θ^* , η^* or β^* in the corresponding model.

¹⁴Note that we account for all input and output slacks of both volume and ratio measures.

¹⁵Because program (12), or its extended analogue (14), may have alternative optimal values $(\hat{\xi}, \hat{\zeta})$, the efficient target defined by formula (13) may not be unique.

The proofs of the above propositions and the other results are given in Appendix C.

Using the statement of technology $T_{\text{VRS}}^{\text{R}}$ by conditions (2), for practical computations program (12) takes on the form:¹⁶

$$\Delta^* = \max \quad \sum_{i=1}^m \xi_i + \sum_{r=1}^s \zeta_r \quad (14.1)$$

$$\text{s.t.} \quad \sum_{j \in J} \lambda_j Y_j^V - \zeta^V \geq \hat{Y}^V, \quad (14.2)$$

$$\sum_{j \in J} \lambda_j X_j^V + \xi^V \leq \hat{X}^V, \quad (14.3)$$

$$\lambda_j \left(Y_j^R - \left[\hat{Y}^R + \zeta^R \right] \right) \geq \mathbf{0}, \quad \forall j \in J, \quad (14.4)$$

$$\lambda_j \left(X_j^R - \left[\hat{X}^R - \xi^R \right] \right) \leq \mathbf{0}, \quad \forall j \in J, \quad (14.5)$$

$$\mathbf{1}^\top \lambda = 1, \quad (14.6)$$

$$\xi = (\xi^V, \xi^R), \zeta = (\zeta^V, \zeta^R) \geq \mathbf{0}, \lambda \geq \mathbf{0}. \quad (14.7)$$

The two nonlinear constraints (14.4) and (14.5) of the above program can be linearized using a similar approach to the method described above. We first state them as the “either-or” conditions

$$\begin{aligned} & \text{either } \lambda_j = 0 \\ & \text{or } \{ \hat{Y}^R + \zeta^R - Y_j^R \leq \mathbf{0} \\ & \text{and } X_j^R - [\hat{X}^R - \xi^R] \leq \mathbf{0} \}, \quad \forall j \in J. \end{aligned}$$

These conditions are restated in a linear form using the variable vector $\delta \in \{0, 1\}^n$:

$$\lambda_j \leq \delta_j, \quad \forall j \in J, \quad (15.1)$$

$$\hat{Y}^R + \zeta^R - Y_j^R \leq L_4(1 - \delta_j), \quad \forall j \in J, \quad (15.2)$$

$$X_j^R - [\hat{X}^R - \xi^R] \leq L_5(1 - \delta_j), \quad \forall j \in J, \quad (15.3)$$

where vectors L_4 and L_5 have sufficiently large components.

Extending the above results to the case of R-CRS is straightforward. Thus, the R-CRS analogue of program (14) is obtained by removing the normalizing condition (14.6). As a result, the linearizing conditions (15) need modifying, because the components of vector λ are no longer bounded above by 1. As shown in Section 4, in this case the inequality (15.1) should be replaced by (11).

6. Alternative calculations of efficient targets

According to formula (13), the efficient target (X^*, Y^*) of DMU (X_o, Y_o) in the R-VRS technology is obtained by adjusting its projection (\hat{X}, \hat{Y}) by the vectors of optimal slacks $\hat{\xi}$

¹⁶The inequality signs in constraints (14.2) and (14.3) can obviously be changed to equalities. However, constraints (14.4) and (14.5) cannot be changed to equalities. It is straightforward to prove that, at optimality, each (scalar) component inequality in (14.4) and (14.5) is satisfied as equality for at least one $j \in J$, where such j is generally different for each component inequality.

and $\hat{\zeta}$ in program (14). A similar definition applies in the case of R-CRS technology. Below we show that the efficient target (X^*, Y^*) of DMU (X_o, Y_o) can alternatively be calculated using an optimal vector λ in program (14) (or its R-CRS analogue). This alternative approach is useful in understanding the notion of PR-inefficiency introduced below, which is further illustrated by a numerical example in Section 10.

First, consider the case of R-VRS. Let $(\hat{\lambda}, \hat{\xi}, \hat{\zeta})$ be an optimal solution to program (14), and let DMU (X^*, Y^*) be the corresponding efficient target of DMU (X_o, Y_o) defined by formula (13). Define the set $J_0 = \{j \in J \mid \hat{\lambda}_j > 0\}$. The next statement shows that the efficient target (X^*, Y^*) can be calculated in an alternative way, using the notion of R-convex combination introduced by Definition 1.

Proposition 3. *In the case of R-VRS, DMU (X^*, Y^*) is the R-convex combination of observed DMUs (X_j, Y_j) , $j \in J_0$, taken with the corresponding weights $\hat{\lambda}_j$.*

Now consider the case of R-CRS. Let $(\hat{\lambda}, \hat{\xi}, \hat{\zeta})$ be an optimal solution to the R-CRS analogue of program (14). (As discussed above, this program is obtained by removing the normalizing condition (14.6).)

As in the case of R-VRS, the efficient target (X^*, Y^*) of DMU (X_o, Y_o) can be defined by formula (13). An alternative approach based on the optimal vector $\hat{\lambda}$ requires the following definition.

Definition 4. DMU (\hat{X}, \hat{Y}) is the ratio-conical (R-conical) combination of DMUs $(\tilde{X}_k, \tilde{Y}_k)$, $k \in \mathcal{K} = \{1, \dots, K\}$, $K \geq 1$, taken with the weights $\lambda_k > 0$, if the four conditions (1) are satisfied.

Note that the only difference between Definition 1 of R-convex combinations and Definition 4 of R-conical combinations is that the former assumes the normalizing condition $\sum_{k \in \mathcal{K}} \lambda_k = 1$ on the weights λ_k , which is omitted in the latter.

It is straightforward to verify that R-conical combinations of DMUs in technology $T_{\text{CRS}}^{\text{F}}$ are also members of this technology. Indeed, in Definition 4, denote $\lambda^* = \sum_{k \in \mathcal{K}} \lambda_k > 0$. The R-conical combination of DMUs $(\tilde{X}_k, \tilde{Y}_k)$ may be viewed as the result of two consecutive actions: taking the R-convex combination of such DMUs with the weights $\hat{\lambda}_k = \lambda_k / \lambda^*$ (which add up to 1), and subsequent proportional scaling of the volume inputs and outputs of the resulting R-convex combination by the factor λ^* , while keeping its ratio measures fixed. Both actions are postulated as feasible in technology $T_{\text{CRS}}^{\text{F}}$, and the resulting DMU is a member of this technology (Olesen et al. 2015).

Proposition 4. *In the case of R-CRS, DMU (X^*, Y^*) is the R-conical combination of observed DMUs (X_j, Y_j) , $j \in J_0$, taken with the corresponding weights $\hat{\lambda}_j$.*

It is worth highlighting the similarity between the above results and the known fact about efficient targets in the standard VRS and CRS models of Banker et al. (1984) and Charnes et al. (1978). In the VRS model, an efficient target of DMU (X_o, Y_o) can be calculated as the convex combination of observed DMUs (X_j, Y_j) , $j \in J$, taken with the corresponding optimal weights $\hat{\lambda}_j$. Because in the R-VRS model ratio inputs and outputs do not form convex combinations, the calculation of efficient targets requires construction of R-convex combinations of observed DMUs.

In the CRS model, the efficient target is obtained by calculating the conical (nonnegative linear) combination of observed DMUs with the optimal vector λ . In the R-CRS model, this is replaced by the R-conical combination of observed DMUs. This excludes ratio inputs and outputs from forming conical combinations and requires taking them at the most conservative level across the combined observed DMUs.

Table 1: Hospitals in Example 1.

Hospital	Costs	Inpatients	Outpatients	Success rate
A	1	1	3	0.8
B	1	3	1	0.9
C	1	2	2	0.8

7. Potential ratio efficiency

Let T^R be any technology with ratio inputs and outputs that satisfies the axioms of selective convexity and free disposability stated in Olesen et al. (2015). As noted above, any such technology necessarily includes all R-convex combinations that can be formed by its DMUs. The R-VRS and R-CRS technologies considered in this paper are two examples of such technology T^R .

While a DMU (\hat{X}, \hat{Y}) may be strongly efficient in technology T^R in the sense of Definition 2, from a practical perspective it might be useful to further explore if some of its ratio inputs or outputs have a potential for improvement if information about the numerators and denominators used for the calculation of ratio measures becomes available. The need for such investigation is highlighted by the following example in which T^R is the R-VRS technology T_{VRS}^R .

Example 1. Consider hospitals A , B and C in Table 1. Their single input represents costs, and the three outputs are inpatients, outpatients and success rate of a certain treatment. The input and the first two outputs are volume measures. (The exact units of measurement are not important for this example.) The last output is the ratio of successful treatments to all treatments undertaken by a hospital, and the underlying data are not available.

Let T_{VRS}^R be the R-VRS technology generated by the three hospitals. It is straightforward to verify that all three hospitals are *strongly* efficient.

Consider hospital C . It is clear that neither A nor B performs better than C . Suppose we want to combine hospitals A and B in a convex combination with the equal weights 0.5. This would not be problematic if all inputs and outputs were volume measures. However, output 3 is a ratio. According to the axiom of selective convexity discussed above, the volume measures are combined in the standard way, while ratio output 3 (success rate) is, to reflect the worst-case scenario, taken at the minimum level 0.8. The resulting aggregated DMU is the R-convex combination of hospitals A and B , which is identical to hospital C .

Although hospital C is strongly efficient in technology T_{VRS}^R and therefore there exists no DMU in this technology that dominates C (even in the weak sense), we can still argue that the ratio output of hospital C could be improved. Indeed, if the volume data about the number of all and successful treatments (instead of their ratio) were available, then the true success rate of hospital C would be strictly between the success rates of hospitals A and B , i.e., strictly between 0.8 and 0.9.

Using the provided information in the ratio form, we cannot estimate the exact improvements on the ratio output 3 available to hospital C . However, we can at least highlight a *potential* for such improvements that in practice may justify further research or a requirement of better (volume) data.

Taking into consideration this argument, we can say that hospital C exhibits *potential ratio (PR) inefficiency*, or that C is PR-inefficient. We define this formally, by using the notion of R-convex combination introduced by Definition 1.

Definition 5. A strongly efficient DMU $(\hat{X}, \hat{Y}) \in T^R$ is *potentially ratio inefficient* (PR-inefficient) if the following two conditions are satisfied:

(i) (\hat{X}, \hat{Y}) is a R-convex combination of a finite number of DMUs $(\tilde{X}_k, \tilde{Y}_k) \in T^R$, $k \in \mathcal{K} = \{1, \dots, K\}$;

(ii) there exists a ratio input $i' \in I^R$ such that $\hat{X}_{i'}^R > \min_{k \in \mathcal{K}} \{\tilde{X}_{ki'}^R\}$, or there exists a ratio output $r' \in O^R$ such that $\hat{Y}_{r'}^R < \max_{k \in \mathcal{K}} \{\tilde{Y}_{kr'}^R\}$.

Otherwise, DMU (\hat{X}, \hat{Y}) is *PR-efficient*.

Note that Definition 5 does not require that the DMUs $(\tilde{X}_k, \tilde{Y}_k)$, $k \in \mathcal{K}$, be observed. The reason for this is that all DMUs in technology T^R are assumed producible and can be used to test for PR inefficiency. If in Example 1 hospitals A and B are not observed, we can still argue that hospital C should be able to increase its success rate, i.e., we still consider it as PR-inefficient.

Also note that the notion of PR inefficiency equally applies to the technology of Ruggiero (1996), which is a special case of the R-VRS technology of Olesen et al. (2015).

8. PR efficiency in the R-VRS technology

8.1. Properties of PR efficiency in the R-VRS technology

Assume that technology T^R is the R-VRS technology T_{VRS}^R . The following result shows that we can verify PR efficiency of DMU (\hat{X}, \hat{Y}) by benchmarking it against R-convex combinations of *observed* DMUs only. This result simplifies computational procedures for testing PR efficiency developed below.

Proposition 5. A strongly efficient DMU $(\hat{X}, \hat{Y}) \in T_{\text{VRS}}^R$ is PR-inefficient if and only if it is a R-convex combination of a finite number of **observed** DMUs $(\tilde{X}_k, \tilde{Y}_k)$, $k \in \mathcal{K} \subseteq J$, for which condition (ii) of Definition 5 is true.

The next result shows that the notion of PR efficiency can in principle be derived from properties of optimal solutions of program (14) used to test for strong efficiency of DMU (\hat{X}, \hat{Y}) .

Proposition 6. A strongly efficient DMU $(\hat{X}, \hat{Y}) \in T_{\text{VRS}}^R$ is PR-efficient if and only if, at each optimal solution to program (14) (whose optimal value in this case is $\Delta^* = 0$), the vector inequalities (14.4) and (14.5) are satisfied as equalities, for all $j \in J$.

8.2. Testing for PR efficiency

Below we develop a test that allows us to identify whether a strongly efficient DMU $(\hat{X}, \hat{Y}) = (\hat{X}^V, \hat{X}^R, \hat{Y}^V, \hat{Y}^R) \in T_{\text{VRS}}^R$ is PR-efficient.¹⁷ Because DMU (\hat{X}, \hat{Y}) is strongly efficient, the optimal value Δ^* in program (14) is zero. This implies that the following

¹⁷DMU (\hat{X}, \hat{Y}) may be observed or unobserved.

program is feasible and has an optimal solution.

$$\Delta_1 = \max \sum_{j \in J} \sum_{i \in I^R} \xi_{ji}^R + \sum_{j \in J} \sum_{r \in O^R} \zeta_{jr}^R \quad (16.1)$$

$$\text{s.t.} \quad \sum_{j \in J} \lambda_j Y_j^V = \hat{Y}^V, \quad (16.2)$$

$$\sum_{j \in J} \lambda_j X_j^V = \hat{X}^V, \quad (16.3)$$

$$\lambda_j \left(Y_j^R - \left[\hat{Y}^R + \zeta_j^R \right] \right) = \mathbf{0}, \quad \forall j \in J, \quad (16.4)$$

$$\lambda_j \left(X_j^R - \left[\hat{X}^R - \xi_j^R \right] \right) = \mathbf{0}, \quad \forall j \in J, \quad (16.5)$$

$$\text{if } \lambda_j = 0 \text{ then } \zeta_j^R = \mathbf{0} \text{ and } \xi_j^R = \mathbf{0}, \quad \forall j \in J, \quad (16.6)$$

$$\mathbf{1}^\top \lambda = 1, \quad (16.7)$$

$$\xi_j^R, \zeta_j^R \geq \mathbf{0}, \quad \forall j \in J, \quad \lambda \geq \mathbf{0}, \quad (16.8)$$

where $\xi_j^R \in \mathbb{R}^{|I^R|}$ and $\zeta_j^R \in \mathbb{R}^{|O^R|}$, $j \in J$, are DMU-specific vectors of slacks.

Let $\hat{\lambda}$, $\hat{\zeta}_j^R$, $\hat{\xi}_j^R$, $j \in J$, be an optimal solution to program (16), and assume that its optimal value $\Delta_1 > 0$. Let J^* be the corresponding reference set of DMU (\hat{X}, \hat{Y}) , i.e., $J^* = \{j \mid \hat{\lambda}_j > 0, j \in J\}$.¹⁸

By constraints (16.4) and (16.5), and because $\hat{\zeta}_j^R, \hat{\xi}_j^R \geq \mathbf{0}$, $\forall j$, none of the DMUs in the reference set J^* performs worse than DMU (\hat{X}, \hat{Y}) on each individual ratio input and output. Moreover, because $\Delta_1 > 0$, there exists a DMU $j \in J^*$ that performs better than DMU (\hat{X}, \hat{Y}) on at least one ratio input or output. Therefore, DMU (\hat{X}, \hat{Y}) exhibits PR inefficiency.

The following result establishes this observation formally.

Proposition 7. *A strongly efficient DMU $(\hat{X}, \hat{Y}) \in T_{\text{VRS}}^R$ is PR-efficient if and only if the optimal value Δ_1 of program (16) is equal to zero.*

It is useful to compare the test of strong efficiency based on program (14) with the test of PR efficiency based on program (16). The former program uses slack vectors ξ^R and ζ^R that are not DMU-specific (these slack vectors do not depend on $j \in J$). As a result, for each ratio measure (input and output) the objective function of program (14) accounts for the *smallest* of all deviations between DMU (\hat{X}, \hat{Y}) and the DMUs from its reference set.¹⁹ In contrast, program (16) uses DMU-specific slack vectors ξ_j^R and ζ_j^R , $j \in J^*$. Therefore, it identifies possible improvements on each ratio measure when DMU (\hat{X}, \hat{Y}) is compared to each observed DMU in the reference set J^* . In summary, only the minima of such improvements enter the objective function of program (14), while *all* improvements enter the objective function of program (16).

Another difference between programs (14) and (16) is that the former program, looking for evidence of weak inefficiency, also accounts for possible slack variables in the volume constraints (14.2) and (14.3). In contrast, in program (16) we assume that DMU (\hat{X}, \hat{Y}) is strongly efficient and, therefore, no slack in the volume inputs and outputs is possible.

¹⁸Program (16) may have multiple optimal solutions, and the reference set J^* may not be unique.

¹⁹This reference set includes all observed DMUs $j \in J$ with a positive λ_j in the optimal solution to program (14).

Therefore, no slack vectors in the volume dimensions are included in the objective function (16.1) and in the constraints (16.2) and (16.3).

In practical applications, program (16) can be restated and solved as a mixed integer linear program. Below we show that a simpler *linear* program can be solved instead.

8.3. An alternative test of PR efficiency

Let, as above, DMU (\hat{X}, \hat{Y}) be strongly efficient. Introduce the set $J_0 \subseteq J$ of all observed DMUs that are not worse than DMU (\hat{X}, \hat{Y}) on all individual ratio inputs and outputs, i.e.,

$$J_0 = \left\{ j \in J \mid X_j^R \leq \hat{X}^R, Y_j^R \geq \hat{Y}^R \right\}.$$

It is clear that DMUs $j \in J \setminus J_0$ cannot enter (with $\lambda_j > 0$) any R-convex combination of observed DMUs that leads to classifying DMU (\hat{X}, \hat{Y}) as PR-inefficient by Definition 5. The following *linear* program uses this fact and restricts the set of observed DMUs, against which DMU (\hat{X}, \hat{Y}) is benchmarked, to J_0 :

$$\Delta_2 = \max \quad \sum_{i \in I^R} \xi_i^R + \sum_{r \in O^R} \zeta_r^R \quad (17.1)$$

$$\text{s.t.} \quad \sum_{j \in J_0} \lambda_j Y_j^V = \hat{Y}^V, \quad (17.2)$$

$$\sum_{j \in J_0} \lambda_j X_j^V = \hat{X}^V, \quad (17.3)$$

$$\sum_{j \in J_0} \lambda_j Y_j^R - \zeta^R = \hat{Y}^R, \quad (17.4)$$

$$\sum_{j \in J_0} \lambda_j X_j^R + \xi^R = \hat{X}^R, \quad (17.5)$$

$$\sum_{j \in J_0} \lambda_j = 1, \quad (17.6)$$

$$\xi^R, \zeta^R \geq \mathbf{0}, \lambda \geq \mathbf{0}. \quad (17.7)$$

Program (17) may be viewed as the additive model based on the VRS technology that treats ratio inputs and outputs in the same way as volume measures, and which accounts only for the slack vectors on the ratio inputs and outputs. Furthermore, this program employs a restricted set J_0 of observed DMUs against which DMU (\hat{X}, \hat{Y}) is benchmarked.

In practical computations we may replace the summation by $j \in J_0$ in constraints (17.2)–(17.6) by the summation by $j \in J$. In this case, to make sure that $\lambda_j = 0$, for all $j \in J \setminus J_0$, we should additionally specify two groups of linear constraints similar to those specified in the statement (2) of technology T_{VRS}^R . These additional constraints take on the form: $\lambda_j (Y_j^R - \hat{Y}^R) \geq \mathbf{0}$ and $\lambda_j (X_j^R - \hat{X}^R) \leq \mathbf{0}$, for all $j \in J$.

Proposition 8. *A strongly efficient DMU $(\hat{X}, \hat{Y}) \in T_{\text{VRS}}^R$ is PR-efficient if and only if the optimal value Δ_2 of program (17) is equal to zero.*

Solving linear program (17) provides a simple practical method for testing PR efficiency of DMUs in technology T_{VRS}^R .

Table 2: Hospitals in Example 2.

Hospital	Costs	Inpatients	Outpatients	Success rate
A^*	0.5	0.5	1.5	0.8
B^*	2	6	2	0.9
C	1	2	2	0.8

9. PR efficiency in the R-CRS technology

The notion of PR efficiency applies to any freely disposable technology that satisfies the assumption of selective convexity. In particular, it applies to the general R-CRS technology of Olesen et al. (2015) that allows for different types of ratio inputs and outputs. However, a full development of testing procedures for PR efficiency in a general R-CRS technology, although straightforward in principle, would require a lengthy technical exposition and, for this reason, is left outside the scope of this paper.

Below we consider such testing procedures only for a special but practically important case of R-CRS technology in which all ratio measures are of the fixed type. The meaning of this technology, also denoted $T_{\text{CRS}}^{\text{F}}$, was discussed in Section 4.

Example 2. Consider a modification of Example 1 in which hospitals A^* , B^* and C have inputs and outputs as in Table 2. (Hospital C is the same as in Example 1.) Suppose that costs, inpatients and outpatients are assumed to be fully scalable (proportional), while the success rate for treatments is fixed and independent of any such scaling.

If the volume input and two outputs of hospitals A^* and B^* are scaled by factors 2 and 0.5, respectively, while keeping the success rate fixed, the resulting hospitals are hospitals A and B in Example 1, although these two hospitals are no longer observed.

As noted in Example 1, hospital C is the R-convex combination of hospitals A and B taken with equal weights 0.5. Because the success rate 0.9 of hospital B is higher than the success rate of 0.8 of hospital C , by Definition 5, the latter hospital is PR-inefficient.

Note that, in contrast with Example 1, the fact that hospital C is PR-inefficient is now established by benchmarking it against *unobserved* DMUs A and B . This is allowed by Definition 5 which, as discussed, does not require that the DMUs used in the R-convex combination be necessarily observed.

Importantly, in contrast with Example 1, hospital C is not a R-convex combination of *observed* hospitals A^* and B^* . Therefore, its PR inefficiency cannot be established if it is benchmarked only against R-convex combinations of observed DMUs.

It is clear that, if we wish to test PR efficiency of any DMU $(\hat{X}, \hat{Y}) \in T_{\text{CRS}}^{\text{F}}$ by benchmarking it against observed DMUs only, we need to allow their R-conical combinations introduced by Definition 4, instead of the more restrictive R-convex combinations. For example, the volume measures of hospital C are obtained as the conical combination of the input and two volume outputs of observed hospitals A^* and B^* taken with the weights $\lambda_A = 1$ and $\lambda_B = 0.25$, while the success rate is kept at the minimum level of 0.8.

Below we generalize this example and state formally that, instead of testing PR efficiency of DMU $(\hat{X}, \hat{Y}) \in T_{\text{CRS}}^{\text{F}}$ by Definition 5, requiring that we consider R-convex combinations of all (observed and unobserved) DMUs in technology $T_{\text{CRS}}^{\text{F}}$, we can benchmark DMU (\hat{X}, \hat{Y}) against R-conical combinations of observed DMUs only.

Let DMU $(\hat{X}, \hat{Y}) \in T_{\text{CRS}}^{\text{F}}$ be strongly efficient in technology $T_{\text{CRS}}^{\text{F}}$, which can be tested

in two stages, as outlined in Sections 4 and 5.²⁰ The following statement is an analogue of Proposition 5.

Proposition 9. *A strongly efficient DMU $(\hat{X}, \hat{Y}) \in T_{\text{CRS}}^{\text{F}}$ is PR-inefficient if and only if it is a R-conical combination of a finite number of **observed** DMUs $(\tilde{X}_k, \tilde{Y}_k)$, $k \in \mathcal{K} \subseteq J$, for which condition (ii) of Definition 5 is true.*

Note that Example 2 provides an illustration to the above statement. The proof of Proposition 9 is similar to the proof of Proposition 5, and is not given.

In practical computations, verifying PR efficiency in technology $T_{\text{CRS}}^{\text{F}}$ is facilitated by the analogues of Propositions 6–8 discussed below.

Restating Propositions 6 and 7 requires removal of the normalizing conditions (14.6) and (16.7) in programs (14) and (16), respectively. The proof of these statements requires only a minor adjustment to the proof in the case of R-VRS, and is not given.

Consider the following modification of program (17):²¹

$$\Delta_3 = \max \quad \sum_{i \in \text{I}^R} \xi_i^R + \sum_{r \in \text{O}^R} \zeta_r^R \quad (18.1)$$

$$\text{s.t.} \quad \sum_{j \in J_0} \lambda_j Y_j^V = \hat{Y}^V, \quad (18.2)$$

$$\sum_{j \in J_0} \lambda_j X_j^V = \hat{X}^V, \quad (18.3)$$

$$\sum_{j \in J_0} \lambda_j (Y_j^R - \hat{Y}^R) - \zeta^R = \mathbf{0}, \quad (18.4)$$

$$\sum_{j \in J_0} \lambda_j (X_j^R - \hat{X}^R) + \xi^R = \mathbf{0}, \quad (18.5)$$

$$\xi^R, \zeta^R \geq \mathbf{0}, \lambda \geq \mathbf{0}. \quad (18.6)$$

Similar to program (17), we can replace condition $j \in J_0$ in constraints (18.2)–(18.5) by the more operational condition $j \in J$, accompanied by the incorporation of additional linear constraints that imply $\lambda_j = 0$, for all $j \in J \setminus J_0$.

The following statement is an analogue of Proposition 8. It provides a straightforward linear programming approach for testing PR efficiency in technology $T_{\text{CRS}}^{\text{F}}$, by solving program (18).

Proposition 10. *A strongly efficient DMU $(\hat{X}, \hat{Y}) \in T_{\text{CRS}}^{\text{F}}$ is PR-efficient if and only if the optimal value Δ_3 of program (18) is equal to zero.*

10. Numerical example and discussion

Below we consider a numerical example that demonstrates different efficiency measures and computational procedures outlined in this paper.

Table 3 shows eight hypothetical observed DMUs evaluated on three inputs and two outputs. In this data set, Inputs 1 and 2, and Output 1 are volume measures, while Input 3 and Output 2 are ratio measures.

Our analysis includes the following three main steps.

²⁰In line with footnote 9, below we assume without further mention that the vector of volume outputs $\hat{Y}^V \neq \mathbf{0}$. This allows us to state technology $T_{\text{CRS}}^{\text{F}}$ in the form (2) from which the normalizing condition $\mathbf{1}^\top \lambda = 1$ is removed.

²¹Program (17) is obtained from program (18) by incorporating the normalizing condition (17.6). Then a simple rearrangement of the terms in constraints (18.4) and (18.5) transforms them to (17.4) and (17.5).

Table 3: The data set.

DMU	Input 1	Input 2	Input 3	Output 1	Output 2
<i>A</i>	15	13	0.2	10	0.7
<i>B</i>	27	25	0.5	16	0.4
<i>C</i>	62	30	0.5	20	0.6
<i>D</i>	25	15	0.4	20	0.5
<i>E</i>	50	35	0.8	16	0.6
<i>F</i>	79	50	0.3	23	0.4
<i>G</i>	25	18	0.6	18	0.3
<i>H</i>	15	25	0.5	20	0.4

1. We evaluate the input radial efficiency of the eight observed DMUs in the R-VRS and R-CRS models, *with respect to the volume inputs only*. For comparison purposes, we also evaluate the efficiency of these DMUs in the free disposal hull (FDH) technology of Deprins et al. (1984), and in the standard CRS and VRS technologies of Charnes et al. (1978) and Banker et al. (1984). Because the input improvement factor θ is attached only to the volume inputs, the R-VRS model becomes the known model of Ruggiero (1996), and the VRS model becomes the model of Banker and Morey (1996) developed for treating exogenously fixed inputs or outputs.
2. Limiting our analysis to the R-VRS and R-CRS models, we test whether the radial efficient DMUs and input radial projections of inefficient DMUs are strongly or only weakly efficient. This illustrates the use of the second optimization stage for the R-VRS and R-CRS models developed above. This step identifies the set of strongly efficient observed DMUs and also efficient targets for the remaining inefficient DMUs.
3. We test whether the strongly efficient observed DMUs and efficient targets (of inefficient DMUs) are PR-efficient, in both the R-VRS and R-CRS models. We find that, in our example, all strongly efficient observed DMUs, and some of the efficient target DMUs, are PR-efficient. The remaining efficient target DMUs are PR-inefficient. We discuss the meaning of these results from a practical perspective.

At the end, we consider an additional scenario in which the input radial efficiency is evaluated with respect to all volume and ratio inputs. Our findings are similar, but not identical, to those discussed within the above main scenario.

10.1. Input radial efficiency in different models

Table 4 shows the input radial efficiency of all DMUs in five different technologies, where the input improvement factor θ is attached only to the volume Inputs 1 and 2.²² This corresponds to a common practical situation in which, for example, ratio inputs and outputs are considered exogenous and nondiscretionary, e.g., reflecting the quality of environment in which the DMUs operate (Banker and Morey 1986, Ruggiero 1996, Haelermans and Ruggiero 2013, Johnson and Ruggiero 2014).

The five models employed in this example are based on different assumptions about the production technology, in particular, about the convexity and scalability of the volume and ratio inputs and outputs.

²²All models in this example were programmed and solved in a common spreadsheet linear optimizer.

Table 4: Input radial efficiency with respect to the two volume inputs in different models.

DMU	FDH	R-VRS	R-CRS	VRS	CRS
<i>A</i>	1	1	1	1	1
<i>B</i>	0.926	0.677	0.615	0.677	0.627
<i>C</i>	1	1	0.867	1	0.541
<i>D</i>	1	1	1	1	1
<i>E</i>	1	0.864	0.594	0.568	0.419
<i>F</i>	1	1	0.598	1	1
<i>G</i>	1	0.874	0.837	0.874	0.837
<i>H</i>	1	1	1	1	1

Free disposal hull (FDH). The FDH technology T_{FDH} (Deprins et al. 1984) does not employ any convexity or scalability (proportionality) assumptions about the inputs and outputs, and is the smallest of the five technologies used in this example. Consequently, the FDH model is the least discriminating among the five models shown in Table 4, in terms of input radial efficiency.

R-VRS technology. The R-VRS technology $T_{\text{VRS}}^{\text{R}}$ (Olesen et al. 2015) is specifically developed for use with a combination of volume and ratio measures, as in the current example. This approach assumes that DMUs can be combined in R-convex combinations. As discussed above, this means that the volume inputs and outputs of the combined DMUs can form the conventional convex combinations, while the ratio inputs and outputs should be taken at their least demanding levels across the combined DMUs.

To evaluate the input radial efficiency in the R-VRS model with respect to the two volume inputs, for each of the eight DMUs, we solve program (3) in which we remove the factor θ from the constraints (3.5). This makes program (3) a linear program. Since the input improvement factor θ is attached only to the volume inputs, the R-VRS model used here is simply the model of Ruggiero (1996).

As seen from Table 4, the R-VRS model provides a better discrimination on efficiency than the FDH model.

R-CRS technology. The R-CRS technology introduced by Olesen et al. (2015) is a generic term for a class of technologies. In this example we employ the practically important R-CRS technology $T_{\text{CRS}}^{\text{F}}$ discussed in Section 4. It assumes that DMUs can be combined in R-convex combinations (as in the R-VRS technology) and that, additionally, volume inputs and outputs can be scaled proportionally with an arbitrary factor $\alpha \geq 0$, while keeping the ratio inputs and outputs fixed. This corresponds to a practical scenario in which the volume of operations can be scaled while keeping exogenous environmental or quality factors (represented by ratio inputs and outputs) fixed.

Assessing the input radial efficiency in technology $T_{\text{CRS}}^{\text{F}}$ was discussed in Section 4. Following this discussion, for each DMU we solve program (3) from which we remove the normalizing equation (3.6). We also remove the input improvement factor θ from the ratio input constraints (3.5), which results in a linear program.

The results of computation in the R-CRS technology $T_{\text{CRS}}^{\text{F}}$ shown in Table 4 illustrate a significant improvement on efficiency discrimination compared to the first two models.

VRS and CRS technologies. Table 4 also shows the input radial efficiency of all DMUs in the standard CRS and VRS technologies T_{CRS} and T_{VRS} of Charnes et al. (1978) and Banker

et al. (1984), evaluated with respect to the volume inputs. Since the input improvement factor θ is attached only to the volume inputs, the VRS model is simply the model proposed by Banker and Morey (1996).

Olesen et al. (2015) argue that the VRS and CRS technologies should not be used if at least one input or output is a ratio measure (and that appropriate substitutes are the R-VRS and R-CRS technologies). This is because both technologies assume the convexity property which is generally violated by the ratio measures. In the case of CRS, a further consideration is that it is generally incorrect to allow proportional scaling of DMUs if their input and output vectors include both volume and ratio measures.

Following these arguments, it may appear meaningless to compare the results of efficiency computations using the R-VRS and R-CRS models, with the results obtained from the VRS and CRS models. However, in our opinion, it is interesting, from a purely mathematical point of view, to illustrate the impact of the different sets of axioms underlying different models, on the resulting efficiency scores. We also find it interesting to explore the mathematical relationship between the different technologies, which is the topic for the next section.

10.2. Relationship between different technologies

The efficiency scores in Table 4 reflect a theoretical relationship between the five different technologies used in our example.

As proved by Proposition 3 in Olesen et al. (2015), the R-VRS technology $T_{\text{VRS}}^{\text{R}}$ is a subset of any R-CRS technology (based on the same set of inputs and outputs), including technology $T_{\text{CRS}}^{\text{F}}$. It is also clear that the FDH technology T_{FDH} is a subset of technology $T_{\text{VRS}}^{\text{R}}$. This can be expressed as follows:

$$T_{\text{FDH}} \subseteq T_{\text{VRS}}^{\text{R}} \subseteq T_{\text{CRS}}^{\text{F}}. \quad (19)$$

It is also straightforward to prove the following embedding:

$$T_{\text{FDH}} \subseteq T_{\text{VRS}}^{\text{R}} \subseteq T_{\text{VRS}} \subseteq T_{\text{CRS}}. \quad (20)$$

Note that the efficiency scores in Table 4 are consistent with the embeddings (19) and (20).

It is worth noting that neither of the two technologies $T_{\text{CRS}}^{\text{F}}$ nor T_{CRS} is a subset of the other.²³ This fact can also be observed from Table 4. Indeed, the input radial efficiency of DMUs B and F in technology $T_{\text{CRS}}^{\text{F}}$ is lower than in technology T_{CRS} , while the opposite is true for DMUs C and E .

10.3. Testing for strong efficiency

Below we illustrate the testing of strong efficiency in the R-VRS and R-CRS models, following the discussion in Sections 5 and 6.

R-VRS model. According to the results in Table 4, five of the eight DMUs, A , C , D , F and H , are input radial efficient in this model (with respect to the two volume inputs only), and three are input radial inefficient. The input radial projections of the three inefficient DMUs, B , E and G , are calculated by multiplying their volume inputs by their input radial efficiency.

²³Indeed, consider technologies $T_{\text{CRS}}^{\text{F}}$ and T_{CRS} generated by the single DMU $W = (1; 0.2; 1; 0.2)$, where the four components, from left to right, represent the volume input, ratio input, volume output and ratio output. Then DMU $W_1 = (2; 0.4; 2; 0.4)$ obtained by full proportional scaling of DMU W by a factor $\alpha = 2$, is in technology T_{CRS} , but not in $T_{\text{CRS}}^{\text{F}}$. However, keeping the ratio input and output of DMU W fixed, and scaling only the volume measures by $\alpha = 2$, we obtain DMU $W_2 = (2; 0.2; 2; 0.2)$. DMU W_2 is in technology $T_{\text{CRS}}^{\text{F}}$ but not in T_{CRS} . Therefore, neither technology $T_{\text{CRS}}^{\text{F}}$ and T_{CRS} is generally a subset of the other.

Table 5: Optimal solutions of the second optimization stage of the R-VRS model: vectors $\lambda = (\lambda_A, \dots, \lambda_H)$, and input and output slack vectors $\xi = (\xi_1, \xi_2, \xi_3)$ and $\zeta = (\zeta_1, \zeta_2)$.

DMU	λ_A	λ_B	λ_C	λ_D	λ_E	λ_F	λ_G	λ_H	ξ_1	ξ_2	ξ_3	ζ_1	ζ_2
<i>A</i>	1	0	0	0	0	0	0	0	0	0	0	0	0
<i>B</i>	0.4	0	0	0.328	0	0	0	0.272	0	0	0	0	0
<i>C</i>	0	0	1	0	0	0	0	0	0	0	0	0	0
<i>D</i>	0	0	0	1	0	0	0	0	0	0	0	0	0
<i>E</i>	0.4	0	0	0.6	0	0	0	0	0	7.04	0.3	0	0
<i>F</i>	0	0	0	0	0	1	0	0	0	0	0	0	0
<i>G</i>	0.2	0	0	0.686	0	0	0	0.114	0	0	0.1	0	0.1
<i>H</i>	0	0	0	0	0	0	0	1	0	0	0	0	0

Table 6: Efficient targets of inefficient DMUs *B*, *E* and *G* in the R-VRS model.

DMU	Input 1	Input 2	Input 3	Output 1	Output 2
<i>B</i> *	18.277	16.923	0.5	16	0.4
<i>E</i> *	43.2	23.2	0.5	16	0.6
<i>G</i> *	21.86	15.74	0.5	18	0.4

To test if the projected DMUs (or the actual DMUs, if these are input radial efficient) are strongly, or only weakly, efficient, we employ program (14). For example, in the case of DMU *A*, which is radial efficient, we replace the DMU under the assessment (\hat{X}, \hat{Y}) in program (14) by DMU *A*. In the case of radial inefficient DMU *B*, we replace (\hat{X}, \hat{Y}) by the projection of *B*, by applying the input improvement factor 0.677 to its two volume inputs.

Solving program (14) requires assessing vectors L_4 and L_5 for conditions (15). Based on the discussion in Appendix B, we take all components of vectors L_4 and L_5 equal to 1.

Table 5 shows the optimal solutions of the second stage R-VRS program (14) for each DMU. The results of the two evaluation stages shown in Tables 4 and 5 can now be combined and summarized as follows.

Five DMUs, *A*, *C*, *D*, *F* and *H*, are strongly efficient in technology $T_{\text{VRS}}^{\text{R}}$. For each of them, the input radial efficiency (with respect to the volume inputs) is equal to 1, and the optimal input and output slacks calculated by the second optimization stage (shown in the last five columns of Table 5) are all equal to zero.

The input radial efficiency of DMU *B* in technology $T_{\text{VRS}}^{\text{R}}$ is equal to 0.677. Its radial projection on the boundary of this technology is obtained by scaling its two volume inputs by 0.677. According to Table 5, this projection cannot be further improved by maximizing the sum of input and output slacks (all optimal slacks for this DMU are equal to zero). Therefore, the input radial projection of DMU *B* is located on the efficient part of the frontier of technology $T_{\text{VRS}}^{\text{R}}$ and is strongly efficient. Consequently, DMU *B* is input radial inefficient but, using the terminology of Cooper et al. (2007), does not exhibit mix inefficiency.

DMUs *E* and *G* are both input radial and mix inefficient. Their radial projections on the boundary of technology $T_{\text{VRS}}^{\text{R}}$, obtained by multiplying their volume inputs by the corresponding efficiency scores 0.864 and 0.874, are only weakly efficient. The latter is observed from the positive optimal input and output slacks shown in Table 5.

Table 7: Optimal solutions of the second optimization stage of the R-CRS model: vectors $\lambda = (\lambda_A, \dots, \lambda_H)$, and input and output slack vectors $\xi = (\xi_1, \xi_2, \xi_3)$ and $\zeta = (\zeta_1, \zeta_2)$.

DMU	λ_A	λ_B	λ_C	λ_D	λ_E	λ_F	λ_G	λ_H	ξ_1	ξ_2	ξ_3	ζ_1	ζ_2
<i>A</i>	1	0	0	0	0	0	0	0	0	0	0	0	0
<i>B</i>	0	0	0	0.462	0	0	0	0.338	0	0	0	0	0
<i>C</i>	2	0	0	0	0	0	0	0	23.733	0	0.3	0	0.1
<i>D</i>	0	0	0	1	0	0	0	0	0	0	0	0	0
<i>E</i>	1.6	0	0	0	0	0	0	0	5.714	0	0.6	0	0.1
<i>F</i>	2.3	0	0	0	0	0	0	0	12.742	0	0.1	0	0.3
<i>G</i>	0	0	0	0.743	0	0	0	0.157	0	0	0.1	0	0.1
<i>H</i>	0	0	0	0	0	0	0	1	0	0	0	0	0

Table 8: Efficient targets of inefficient DMUs *B*, *C*, *E*, *F* and *G* in the R-CRS model.

DMU	Input 1	Input 2	Input 3	Output 1	Output 2
<i>B</i> *	16.615	15.385	0.5	16	0.4
<i>C</i> *	30	26	0.2	20	0.7
<i>E</i> *	24	20.8	0.2	16	0.7
<i>F</i> *	34.5	29.9	0.2	23	0.7
<i>G</i> *	20.93	15.07	0.5	18	0.4

Table 6 shows the efficient targets B^* , E^* and G^* , for the three inefficient DMUs, B , E and G , calculated using the results shown in Tables 4 and 5. According to Proposition 3, we can calculate the efficient targets in two different ways, but the result is always the same.

For example, consider DMU E . Its efficient target E^* shown in Table 6 is obtained as the R-convex combination of the reference DMUs A and D taken with the weights $\lambda_A = 0.4$ and $\lambda_D = 0.6$ shown in Table 5. Alternatively, we obtain the same target DMU E^* by formula (13), i.e., by first calculating the input radial projection of DMU E on the boundary of the technology, and subsequently adjusting this target by the optimal slacks $\xi_2 = 7.04$ and $\xi_3 = 0.3$ from Table 5.

R-CRS model. In the case of R-CRS technology $T_{\text{CRS}}^{\text{F}}$, we use the same program (14) from which we remove the normalizing condition (14.6). We further specify the same vectors L_4 and L_5 for the linearized conditions (15.2) and (15.3), as in the R-VRS model. We also replace condition (15.1) by condition (11) in which we take all components of vector Λ equal to the sufficiently large and “round” value 10.

Table 7 shows the optimal solutions of the second stage R-CRS program. Based on the results in both Tables 4 and 7, DMUs A , D and H are strongly efficient in technology $T_{\text{CRS}}^{\text{F}}$. DMU B is input radial inefficient and does not exhibit mix inefficiency (all optimal slacks in the second stage solution for this DMU, as shown in Table 7, are equal to zero). The remaining four DMUs C , E , F and G are both input radial inefficient and mix inefficient.

Table 8 shows the corresponding efficient targets for the inefficient DMUs. As in the case of R-VRS, these efficient targets can be calculated in two equivalent ways, i.e., by formula (13) or, according to Proposition 4, by taking the R-conical combinations of observed DMUs with the corresponding optimal weights λ .

Table 9: Testing PR efficiency of the efficient target DMUs.

Model	A	B	C	D	E	F	G	H
Δ_2 in R-VRS program (17)	0	0.306	0	0	0.16	0	0.257	0
Δ_3 in R-CRS program (18)	0	0.092	0	0	0	0	0.149	0

10.4. Testing for PR efficiency

Conventional applications of DEA typically involve computations in two optimization stages, leading to the identification of efficient targets of the inefficient DMUs. Because the efficient targets are strongly efficient, they are often regarded as the ultimate goals for the inefficient DMUs seeking to improve their operations.

In this paper we argued that, in technologies $T_{\text{VRS}}^{\text{R}}$ and $T_{\text{CRS}}^{\text{F}}$, incorporating ratio measures that in practical scenarios would often represent environmental or quality factors, there may still be evidence that a strongly efficient DMU could be improved further. This leads to the development of the notion of PR inefficiency that applies only to strongly efficient DMUs. If a strongly efficient DMU is PR-inefficient, it will be considered inefficient if the volume numerators and denominators of the ratio inputs and outputs become available.

Below we show that the notion of PR efficiency is useful for further analysis of efficient targets of all DMUs, in both R-VRS and R-CRS models.

R-VRS model. We test PR efficiency of the five strongly efficient DMUs, A , C , D , F and H , and the efficient targets B^* , E^* and G^* of the inefficient DMUs, by solving appropriately specified linear programs (17). The optimal vectors λ for each of the eight tested DMUs are the same as obtained in the second optimization stage and shown in Table 5, and are not reproduced.²⁴ However, the optimal input and output slacks are generally different (because of their different use in the two models). The optimal sum of these slacks is the value Δ_2 shown in Table 9.

According to these results, by Proposition 8, all five strongly efficient DMUs are also PR-efficient, and the three efficient targets of the inefficient DMUs are PR-inefficient.

Let us explore in detail the meaning of PR inefficiency. Consider, for example, DMU B . Its efficient target is DMU B^* shown in Table 6. According to Proposition 3, and using the optimal solution shown in Table 5, DMU B^* is obtained by taking the R-convex combination of the strongly efficient DMUs A , D and H with the weights $\lambda_A = 0.4$, $\lambda_D = 0.328$ and $\lambda_H = 0.272$, respectively. This fact is also straightforward to verify by direct computations confirming that DMU B^* indeed satisfies all conditions of Definition 1:

- The two volume inputs and the volume output of DMU B^* are equal to the weighted sums of these inputs and outputs of the DMUs A , D and H taken with the optimal weights λ_A , λ_D and λ_H .
- The ratio input of DMU B^* is equal to 0.5, which is the maximum of the ratio inputs 0.2, 0.4 and 0.5 of DMUs A , D and H .
- The ratio output of DMU B^* is equal to 0.4, which is the minimum of the ratio outputs 0.7, 0.5 and 0.4 of DMUs A , D and H .

²⁴If the second-stage program (14) has alternative optimal vectors λ , it is theoretically possible that some of them may not be optimal in program (17) for testing PR efficiency.

As argued in Olesen et al. (2015), the ratio input and output of the target DMU B^* represent the most conservative estimates that have to be assumed because the underlying volume data (used in the numerator and denominator of the ratios) are not available. Repeating the argument used in Example 1, observe that if the volume data become available, and the true ratio inputs and outputs are calculated (and not conservatively assessed as in the R-VRS model), the ratio input of DMU B^* will be strictly smaller than 0.5. More precisely, it will be in the range (0.2; 0.5) defined by the minimum and maximum values of this ratio input among DMUs A , D and H . Similarly, the ratio output of DMU B^* will be strictly larger than 0.4 and, more more precisely, be located in the range (0.4; 0.7).

Based on the available ratio data, the efficient target DMU B^* is strongly efficient and represents an utmost feasible goal for improvement for the inefficient DMU B . However, if the volume measures become available, the ratio input 0.5 and output 0.4 of DMU B^* would no longer represent the best feasible benchmarks for DMU B .

R-CRS model. We use program (18) to test PR efficiency of the three strongly efficient observed DMUs, A , D and H , and the efficient targets B^* , C^* , E^* , F^* and G^* of the five inefficient DMUs, as shown in Table 8. As in the case of R-VRS, in our example, the optimal vectors λ of program (18) are the same as those obtained in the second optimization stage and shown in Table 7.

The optimal values of program (18) calculated for each DMU and denoted Δ_3 are shown in Table 9. By Proposition 10, all strongly efficient observed DMUs, A , D and H , and the efficient targets C^* , E^* and F^* are PR-efficient. The efficient targets B^* and G^* are PR-inefficient.

To explain the meaning of PR-inefficiency in the R-CRS model, we again refer to DMU B as an example. Its efficient target B^* is shown in Table 8. By Proposition 4, and using the optimal vector λ shown in Table 7, DMU B^* is the R-conical combination of the two efficient DMUs D and H , taken with the weights $\lambda_D = 0.462$ and $\lambda_H = 0.338$, respectively. In line with Definition 4, this implies the following:

- The volume inputs and output of DMU B^* are equal to the weighted sums of these inputs and outputs of DMUs D and H taken with the weights λ_D and λ_H .
- The ratio input of DMU B^* is equal to 0.5, which is the maximum of the ratio inputs 0.4 and 0.5 of DMUs D and H .
- The ratio output of DMU B^* is equal to 0.4, which is the minimum of the ratio outputs 0.5 and 0.4 of DMUs D and H .

As in the case of R-VRS, the efficient target B^* represents an utmost feasible goal for improving the inefficient DMU B , based on the available data. However, if the underlying volume data used in the calculation of ratio inputs and outputs become available, the correctly calculated (and not based on the most conservative assumption used in the R-CRS model) ratio input and output of DMU B^* will both be strictly in the range (0.4; 0.5), spanning the minimum and maximum values of the ratio input and output of DMUs D and H . Therefore, the current efficient target B^* will no longer be regarded as a benchmark for the inefficient DMU B .

It is also interesting to comment on the fact that the efficient targets of DMUs C , E and F are PR-efficient. This is because each of these DMUs has a single reference DMU identified by the positive component of the optimal vector λ in program (18), as shown in Table 7.

Table 10: Input radial efficiency with respect to all inputs in different models.

DMU	FDH	R-VRS	R-CRS	VRS	CRS
<i>A</i>	1	1	1	1	1
<i>B</i>	0.926	0.8	0.8	0.688	0.677
<i>C</i>	1	1	0.867	1	0.744
<i>D</i>	1	1	1	1	1
<i>E</i>	1	0.864	0.594	0.568	0.419
<i>F</i>	1	1	0.667	1	1
<i>G</i>	1	0.874	0.837	0.874	0.837
<i>H</i>	1	1	1	1	1

For example, for DMU *C*, the only positive component is $\lambda_A = 2$. Therefore, its efficient target C^* (shown in Table 8) is obtained by multiplying the volume inputs and output of DMU *A* by a factor 2, while keeping the ratio input and output of DMU *A* fixed. Therefore, DMU C^* is the trivial R-conical combination of the single DMU *A*, which does not generate a range of potential ratio values as in the cases discussed above, in which several components of vector λ were positive.²⁵

10.5. Input radial efficiency with respect to all inputs

In this scenario we attach the input improvement factor θ to all three (volume and ratio) inputs, and follow the same three steps as above. In the first step, we assess the input radial efficiency of each DMU in the five different technologies used in the main example above.

We evaluate the efficiency of DMUs in technology $T_{\text{VRS}}^{\text{R}}$ by solving an appropriately specified program (3), in which the nonlinear inequalities (3.5) are replaced by inequalities (5). Similarly, in the case of technology $T_{\text{CRS}}^{\text{F}}$, we solve program (3), in which the normalizing equality $\mathbf{1}^\top \lambda = 1$ is omitted, and the inequality (5.1) is replaced by (11) in which all components of vector Λ are taken equal to 10.²⁶

It is clear that attaching the radial input improvement factor θ to all inputs would generally result in the same or higher efficiency scores compared to our main example in which θ is attached only to a subset of inputs. This is confirmed by the computational results shown in Table 10.

As in the main example above, we focus next on the R-VRS and R-CRS models. The second optimization stage testing for strong efficiency of the input radial projections, and the third stage testing for PR efficiency of the efficient targets do not require any modification and are performed exactly as above.

In the R-VRS model, DMUs *B*, *E* and *G* are input radial inefficient. Performing the second optimization stage shows that the radial projections of these three DMUs also exhibit mix inefficiency.²⁷ The remaining five DMUs, *A*, *C*, *D*, *F* and *H*, are strongly efficient. We test these five DMUs and the efficient targets of the three inefficient DMUs for PR efficiency (the efficient targets are not shown). Computations show that all five strongly efficient DMUs are PR-efficient, while the efficient targets of DMUs *B*, *E* and *G* are PR-inefficient.

²⁵By Proposition 9, the R-conical combination of DMUs in the set \mathcal{K} must satisfy condition (ii) of Definition 5, which is obviously impossible if \mathcal{K} consists of a single DMU.

²⁶This is the same vector Λ as used in our main example for testing strong efficiency in the R-CRS model.

²⁷The optimal sum of slacks in the appropriately specified program (14) is strictly positive.

In the R-CRS model, only three DMUs, A , D and H , are input radial efficient. The second optimization stage shows that these DMUs are strongly efficient. The radial projections of all five input radial inefficient DMUs B , C , E , F and G , exhibit mix inefficiency and are therefore only weakly efficient. Further testing shows that, in the R-CRS model, the three strongly efficient DMUs and the efficient targets of all inefficient DMUs, except DMU G , are PR-efficient. The efficient target of DMU G is PR-inefficient.

Interpretation of PR inefficiency in the R-VRS and R-CRS models is similar to our main scenario and is not given.

11. Conclusion

In this paper we consider practical approaches to solving DEA models with ratio inputs and outputs recently introduced by Olesen et al. (2015). The computational complexity of these models depends on several factors:

- the assumption of returns to scale, i.e., whether we use the ratio variable or constant returns-to-scale (R-VRS or R-CRS) technology;
- the set of inputs and outputs with respect to which the efficiency is evaluated, e.g., whether the improvements involve only volume measures or both volume and ratio measures;
- the specific efficiency measure used and the orientation of the model, e.g., whether we employ an input or output radial efficiency measure, or a nonradial efficiency measure based on a specified direction.

If a radial or nonradial efficiency measure involves assessing improvements only with respect to volume inputs or outputs, the R-VRS and R-CRS DEA models are straightforward linear programs. However, if the efficiency measure includes ratio inputs or outputs, the R-VRS and R-CRS models become nonlinear.

In this paper we show how all nonlinear R-VRS models and a practically important class of R-CRS models with fixed (e.g., environmental) ratio inputs and outputs could be linearized by transformation to a mixed integer linear program. This transformation offers obvious computational advantages in practical applications.

We also consider the second optimization stage seeking to determine if a DMU is strongly or only weakly efficient. This is achieved by maximising the sum of input and output slacks, similar to the standard DEA models. The new second-stage models with ratio inputs and outputs are nonlinear but can be linearized in a way similar to our main development.

As in the case of standard VRS and CRS models of Banker et al. (1984) and Charnes et al. (1978), we show that the second optimization stage for the R-VRS and R-CRS models leads to the identification of efficient targets of inefficient DMUs. However, while in the standard VRS and CRS models the efficient targets are equal to the convex and, respectively, conical (nonnegative linear) combinations of reference DMUs, in the R-VRS and R-CRS models these targets are calculated as the ratio-convex (R-convex) and, respectively, ratio-conical (R-conical) combinations of the reference DMUs. The notions of R-convex and R-conical combinations are defined in our paper and reflect the distinctly different treatment of the volume and ratio measures when constructing composite DMUs.

In a new conceptual development, we also introduce the notion of potential ratio (PR) efficiency in a arbitrary technology with ratio measures that satisfies the axiom of selective convexity and is freely disposable in all inputs and outputs. This notion applies to strongly

efficient DMUs that can be shown to become inefficient if the volume data used to calculate ratio measures become available. We develop computational approaches for testing PR efficiency in the R-VRS technology and in the practically important R-CRS technology with fixed ratio inputs and outputs.

We conclude our paper by a numerical example that illustrates different efficiency concepts (radial efficiency, strong and weak efficiency, PR efficiency) and computational approaches for their evaluation. In particular, we show that the notion of PR efficiency is useful for analysis of efficient target DMUs in both the R-VRS and R-CRS models. If the volume data used in the calculation of ratio inputs and outputs become available, the ratio inputs and outputs of all PR-inefficient target DMUs will become more demanding and located within the particular ranges that are straightforward to calculate.

In practical terms this means that PR-inefficient target DMUs may be considered only as intermediate benchmarks for the inefficient DMUs, or the *most conservative estimates* of the “true” benchmarks, that we use in the R-VRS and R-CRS models because we do not have access to the volume data underlying the available ratios. These intermediate benchmarks will no longer represent the best feasible goals for ratio inputs and outputs if the volume data become available.

Appendix A. Computational problems related to the use of big M in linear programming

The known linearization methods of “either-or” conditions require that the vectors L_k , $k = 1, \dots, 5$, employed in this paper, be sufficiently large. In Appendix B we estimate lower bounds for these vectors that satisfy theoretical requirements for the linearization methods. Unfortunately, for some data sets, choosing vectors L_k at or above these bounds, as required by the theory, may lead to well-known computational problems.

As illustrated in Sections 1–4, the large vectors L_k are used to model logical conditions of the type: either condition A must hold or condition B must hold. The vector inequalities containing vectors L_k , such as (5.2), consist of similar *scalar* inequalities that can be stated in the following generic form:

$$f(x) \leq Mb, \tag{A.1}$$

where M is a component of the “large” vector L_k provided by the analyst, $b \in \{0, 1\}$ is a binary variable, and $f(x)$ is a generic scalar function of some variable x .

Because b is binary, the single inequality (A.1) is theoretically equivalent to the following:

$$\begin{aligned} \text{either i) } & b = 0 \Rightarrow f(x) \leq 0 \\ \text{or ii) } & b = 1 \Rightarrow f(x) \leq M. \end{aligned}$$

The idea of choosing a large constant M is to make inequality $f(x) \leq M$ redundant, if $b = 1$.

The generic *numerical instability* of inequality (A.1) arises from the fact that the optimization software uses a floating point representation of the binary variable b . Hence, b is *not* necessarily 0 or 1. Assume that the integer tolerance of the solver is $1e^{-5}$, and that $b = 0.5e^{-5}$ in a current basic feasible solution to a linear program.²⁸ Then b is regarded by the software as being zero. Hence, no further branching or iterating procedure will occur to try to bring b closer to zero. If the user specifies an M larger than the inverse integer tolerance, e.g., if $M = 1e^6$, then the “either-or” constraint modelled by (A.1) may not work properly. The solver will regard $b = 0.5e^{-5}$ as a zero, but (A.1) no longer implies that $f(x) \leq 0$. Indeed, in this case we obtain $f(x) \leq 10^6 \times 0.5e^{-5} = 5 \neq 0$.

Note that the software used to solve a linear program with constraint (A.1) will not provide any warning message if the specified constant M is too large.

Appendix B. Estimation of large vectors L

Below we suggest some approaches to defining “large” vectors from L_1 to L_5 used for the linearization of different models discussed above.

Vector L_1 . The constant vector L_1 is used in condition (5.2). Its role is to make this condition redundant for any $j \in J$ such that $\delta_j = 0$ and, therefore, $\lambda_j = 0$. Define vector \tilde{X}^R of dimension $|I^R|$ as follows:

$$\tilde{X}_i^R = \max_{j \in J} X_{ji}^R, \quad \forall i \in I^R. \tag{B.1}$$

We can now take L_1 as any vector that satisfies the inequality

$$L_1 \geq \tilde{X}^R.$$

²⁸The default integer tolerance of industry standard CPLEX solver is currently $1e^{-5}$.

Suppose we have replaced the nonlinear constraints (3.5) by linear conditions (5). By (5.1), if $\lambda_j > 0$ then $\delta_j = 1$, and (5.2) implies $X_j^R - \theta X_o^R \leq \mathbf{0}$, which is consistent with (3.5). If $\lambda_j = 0$ then $\delta_j = 0$ becomes feasible and (5.2) becomes $X_j^R - \theta X_o^R \leq L_1$. By the definition of vector L_1 , this inequality is true for any value of $\theta \geq 0$. Therefore, if $\lambda_j = 0$, constraints (5) are redundant, exactly as the constraint (3.5).

Vector L_2 . Let \mathcal{P}_1 be program (6) whose nonlinear constraints (6.4) are replaced by two inequalities (7). The components of vector L_2 should be so large that, for every j such that $\delta_j = 0$ (possible only if the corresponding $\lambda_j = 0$), the inequality (7.2) becomes redundant and does not limit the feasible set of program \mathcal{P}_1 . We achieve this by requiring that the inequality $\eta Y_o^R - Y_j^R \leq L_2$ be true for all $j \in J$ and for all feasible solutions of program \mathcal{P}_1 . Because vectors Y_j^R are nonnegative, it suffices to require that $\eta Y_o^R \leq L_2$, for all feasible values η .

Let $M \geq 1$ be an upper bound on η , i.e., $\eta \leq M$ for all feasible values η in program \mathcal{P}_1 . Then L_2 can be defined as any vector that satisfies the following inequality:

$$L_2 \geq M Y_o^R.$$

In program (6) we consider radial expansion of both volume and ratio outputs. This allows us to obtain two independent upper bounds on η . First, define vector $\tilde{Y} \in \mathbb{R}_+^s$ as follows:

$$\tilde{Y}_r = \max_{j \in J} Y_{jr}, \quad \forall r = 1, \dots, s. \quad (\text{B.2})$$

In the case of volume outputs, the inequalities (6.2) imply $\tilde{Y}^V \geq \eta Y_o^V$. Assuming that at least for one output r we have $Y_{or}^V \neq 0$, for any feasible η we have $\eta \leq M_1$, where

$$M_1 = \min\{\tilde{Y}_r^V / Y_{or}^V \mid r \in \text{O}^V : Y_{or}^V \neq 0\}.$$

Considering the ratio outputs, we can derive an upper bound on η from inequalities (6.4). Indeed, in program \mathcal{P}_1 , there exists a $j \in J$ such that $\lambda_j > 0$. Therefore, $\eta Y_o^R - Y_j^R \leq \mathbf{0}$. This implies $\eta \leq M_2^{(j)}$, where $M_2^{(j)} = \min\{Y_{jr}^R / Y_{or}^R \mid r \in \text{O}^R : Y_{or}^R \neq 0\}$, provided $Y_{or}^R \neq 0$ for at least one r . Because j is unknown a priori, we use the conservative upper bound

$$M_2 = \max_{j \in J} M_2^{(j)}.$$

Because from a computational point of view it is important to have L_2 as small as possible (see Appendix A), we can choose $M = \min\{M_1, M_2\}$ and $L_2 = M Y_o^R$. On the other hand, if the model in use considers radial expansion of only volume outputs (or only ratio outputs), we have to accept the possibly larger $M = M_1$ (respectively, $M = M_2$).

Finally, the above constants M , M_1 and M_2 are specific to the DMU (X_o, Y_o) under the evaluation. In practice we may prefer to assess the efficiency of all observed DMUs $j \in J$ using the same large constant M^* . In this case we may first assess an appropriate constant $M(j)$ for each observed DMU j , and then define $M^* = \max_{j \in J} \{M(j)\}$.

Vector L_3 . Let \mathcal{P}_2 be program (9) whose nonlinear constraints (9.5) are replaced by inequalities (10). In program \mathcal{P}_2 , there exists a $j^* \in J$ such that $\lambda_{j^*} > 0$ and hence, by (10.1), $\delta_{j^*} = 1$. Then from (10.2), we have

$$X_o^R - \beta g_X^R \geq X_{j^*}^R \geq \mathbf{0}, \quad (\text{B.3})$$

for all values β feasible in program \mathcal{P}_2 .

Consider any $j \in J$. If $\lambda_j > 0$, then by (10.1), $\delta_j = 1$ and the inequality (10.2) implies $X_j^R - (X_o^R - \beta g_X^R) \leq \mathbf{0}$, as required. If $\lambda_j = 0$ then $\delta_j = 0$ becomes feasible and the corresponding inequality (10.2) takes on the form

$$X_j^R - (X_o^R - \beta g_X^R) \leq L_3. \quad (\text{B.4})$$

We need to define vector L_3 sufficiently large so that the inequality (B.4) is redundant for all $j \in J$ and all β feasible in program \mathcal{P}_2 . Because, by (B.3), $X_o^R - \beta g_X^R \geq \mathbf{0}$, to satisfy (B.4), it suffices to define L_3 as any vector such that

$$L_3 \geq \tilde{X}^R,$$

where the vector \tilde{X}^R is as defined in (B.1). Therefore L_3 can be the same vector as L_1 .

Vectors L_4 and L_5 . The components of vectors L_4 and L_5 should be sufficiently large to render inequalities (15.2) and (15.3) redundant for all j such that $\delta_j = 0$. Let \mathcal{P}_3 denote program (14) whose nonlinear constraints (14.4) and (14.5) are replaced by the three inequalities (15).

Note that in program \mathcal{P}_3 , there exists a $j^* \in J$ such that $\lambda_{j^*} > 0$ and therefore, by (15.1), $\delta_{j^*} = 1$. Then (15.2) implies

$$\hat{Y}^R + \zeta^R \leq Y_{j^*}^R \leq \tilde{Y}^R, \quad (\text{B.5})$$

where \tilde{Y}^R is the subvector of ratio outputs of the vector \tilde{Y} defined by (B.2). For any feasible solution of \mathcal{P}_3 and for any $j \in J$, for the left-hand side of inequality (15.2) we have:

$$\hat{Y}^R + \zeta^R - Y_j^R \leq \tilde{Y}^R - Y_j^R \leq \tilde{Y}^R,$$

where the left inequality follows from (B.5). Define L_4 as any vector that satisfies the inequality

$$L_4 \geq \tilde{Y}^R.$$

Then, for any $j \in J$ such that $\lambda_j = 0$, the value $\delta_j = 0$ is feasible. The corresponding inequality (15.2), which takes on the form

$$\hat{Y}^R + \zeta^R - Y_j^R \leq L_4,$$

is redundant in program \mathcal{P}_3 , as required.

Consider inequalities (15.3). As noted, there exists a j^* such that $\delta_{j^*} = 1$. Then (15.3) implies

$$\hat{X}^R - \xi^R \geq X_{j^*}^R \geq \mathbf{0}. \quad (\text{B.6})$$

Let us show that we can take any vector L_5 such that

$$L_5 \geq \tilde{X}^R,$$

where vector \tilde{X}^R is as defined by (B.1). (We can therefore take $L_5 = L_1$, where L_1 is defined above.) Indeed, for any $j \in J$ such that $\lambda_j = 0$, the value $\delta_j = 0$ is feasible and, for the left-hand side of inequality (15.3), we have

$$X_j^R - [\hat{X}^R - \xi^R] \leq X_j^R \leq \tilde{X}^R \leq L_5,$$

where the left inequality follows from (B.6). Therefore, for any j such that $\delta_j = 0$, the inequality (15.3) is true for all feasible solutions to program \mathcal{P}_3 and is therefore redundant, as required.

Appendix C. Proofs

Proof of Proposition 1. By the statement of program (12), DMU $(X^*, Y^*) \in T_{\text{VRS}}^{\text{R}}$. Assume that (X^*, Y^*) is inefficient. Then there exists a DMU $(X', Y') \in T_{\text{VRS}}^{\text{R}}$ such that $(X', Y') = (X^* - \xi', Y^* + \zeta')$, where at least one of the two vectors $\xi' \in \mathbb{R}_+^m$ and $\zeta' \in \mathbb{R}_+^s$ is not a zero vector. Taking into account (13), we have

$$(\hat{X} - \hat{\xi} - \xi', \hat{Y} + \hat{\zeta} + \zeta') \in T_{\text{VRS}}^{\text{R}}. \quad (\text{C.1})$$

Denote $\tilde{\xi} = \hat{\xi} + \xi'$ and $\tilde{\zeta} = \hat{\zeta} + \zeta'$. Because at least one component of vectors ξ' and ζ' is strictly positive, we have

$$\sum_{i=1}^m \tilde{\xi}_i + \sum_{r=1}^s \tilde{\zeta}_r > \sum_{i=1}^m \hat{\xi}_i + \sum_{r=1}^s \hat{\zeta}_r.$$

The latter inequality, together with condition (C.1), implies that $(\hat{\xi}, \hat{\zeta})$ is not an optimal solution to program (12). Therefore, DMU (X^*, Y^*) is efficient. \square

Proof of Proposition 2. If either of the two conditions (i) or (ii) is not true, then DMU (X_o, Y_o) is obviously dominated and inefficient. Conversely, if (i) is true then $(X_o, Y_o) = (\hat{X}, \hat{Y})$. If (ii) is true then $(\hat{X}, \hat{Y}) = (X^*, Y^*)$, and therefore $(X_o, Y_o) = (X^*, Y^*)$. By Proposition 1, (X_o, Y_o) is efficient. \square

Proofs of Propositions 3 and 4 . Let $(\hat{\lambda}, \hat{\xi}, \hat{\zeta})$ be an optimal solution to program (14), and let (X^*, Y^*) be defined by (13). Then equalities (1.1) and (1.2) are true. If for at least one $r \in O^R$ or $i \in I^R$ the equalities (1.3) and (1.4) are not satisfied, then the corresponding slacks ζ_r^R and ξ_i^R can be increased, and the solution $(\hat{\lambda}, \hat{\xi}, \hat{\zeta})$ is not optimal.

The proof of Proposition 4 is similar and is omitted. \square

Proof of Proposition 5. Let DMU (\hat{X}, \hat{Y}) be PR-inefficient. By Definition 5, (\hat{X}, \hat{Y}) is the R-convex combination of a finite number of DMUs $(\tilde{X}_k, \tilde{Y}_k) \in T_{\text{VRS}}^{\text{R}}$, $k \in \mathcal{K}$, taken with weights λ_k , such that condition (ii) of Definition 5 is true. Each DMU $(\tilde{X}_k, \tilde{Y}_k)$ satisfies conditions (2) with some vector $\mu^k \in \mathbb{R}^n$. (We use notation μ^k to avoid confusion with the weights λ_k .) Therefore, for each $k \in \mathcal{K}$, we have

$$\sum_{j \in J} \mu_j^k Y_j^V \geq \tilde{Y}_k^V, \quad (\text{C.2.1})$$

$$\sum_{j \in J} \mu_j^k X_j^V \leq \tilde{X}_k^V, \quad (\text{C.2.2})$$

$$\mu_j^k (Y_j^R - \tilde{Y}_k^R) \geq 0, \quad \forall j \in J, \quad (\text{C.2.3})$$

$$\mu_j^k (X_j^R - \tilde{X}_k^R) \leq 0, \quad \forall j \in J, \quad (\text{C.2.4})$$

$$\mathbf{1}^\top \mu^k = 1, \quad \mu^k \geq 0, \quad (\text{C.2.5})$$

where (X_j, Y_j) , $j \in J$, are the observed DMUs.

Define vector $\bar{\lambda} \in \mathbb{R}_+^n$ as follows: $\bar{\lambda}_j = \sum_{k \in \mathcal{K}} \lambda_k \mu_j^k$, $\forall j \in J$. Then $\mathbf{1}^\top \bar{\lambda} = 1$ and $\bar{\lambda} \geq 0$. For each $k \in \mathcal{K}$, denote $J_k = \{j \in J \mid \mu_j^k > 0\}$. Let $j \in J^* = \bigcup_{k \in \mathcal{K}} J_k$. Then $\bar{\lambda}_j > 0$ if and only if $j \in J^*$.

The idea of the proof is to show that DMU (\hat{X}, \hat{Y}) is the R-convex combination (\bar{X}, \bar{Y}) of observed DMUs (X_j, Y_j) , $j \in J^*$, taken with the weights $\bar{\lambda}_j$. The proof is finalized by showing that this R-convex combination also satisfies condition (ii) of Definition 5. Components of the R-convex combination $(\bar{X}, \bar{Y}) = (\bar{X}^V, \bar{X}^R, \bar{Y}^V, \bar{Y}^R)$ are calculated below.

For each $k \in \mathcal{K}$, multiply constraints (C.2.1) and (C.2.2) by $\lambda_k > 0$. Summing by k , restricting j to the set J^* , and noting condition (i) in Definition 5, we have

$$\begin{aligned}\bar{Y}^V &= \sum_{j \in J^*} \bar{\lambda}_j Y_j^V \geq \sum_{k \in \mathcal{K}} \lambda_k \tilde{Y}_j^V = \hat{Y}^V, \\ \bar{X}^V &= \sum_{j \in J^*} \bar{\lambda}_j X_j^V \leq \sum_{k \in \mathcal{K}} \lambda_k \tilde{X}_k^V = \hat{X}^V.\end{aligned}\tag{C.3}$$

Let $j \in J^*$. Because $\bar{\lambda}_j > 0$, there exists a $k \in \mathcal{K}$ such that $\mu_j^k > 0$. By (C.2.3), we have $Y_{jr}^R \geq \tilde{Y}_{kr}^R, \forall r \in \mathcal{O}^R$. By (1.3), $\tilde{Y}_{kr}^R \geq \hat{Y}_r^R$. Therefore, $Y_{jr}^R \geq \hat{Y}_r^R$. Similarly, $X_{ji}^R \leq \hat{X}_i^R, \forall i \in \mathcal{I}^R$. Then, by definition of \bar{Y}^R and \bar{X}^R , we have

$$\begin{aligned}\bar{Y}_r^R &= \min_{j \in J^*} \{Y_{jr}^R\} \geq \hat{Y}_r^R, \quad \forall r \in \mathcal{O}^R. \\ \bar{X}_i^R &= \max_{j \in J^*} \{X_{ji}^R\} \leq \hat{X}_i^R, \quad \forall i \in \mathcal{I}^R.\end{aligned}\tag{C.4}$$

By Definition 1, DMU $(\bar{X}, \bar{Y}) = (\bar{X}^V, \bar{X}^R, \bar{Y}^V, \bar{Y}^R)$, whose components are calculated in (C.3) and (C.4), is the R-convex combination of observed DMUs $j \in J^*$ taken with weights $\bar{\lambda}_j > 0$. Because DMU (\hat{X}, \hat{Y}) is strongly efficient, all non-strict inequalities in (C.3) and (C.4) are satisfied as equalities, as otherwise it would be dominated by DMU (\bar{X}, \bar{Y}) . Therefore, $(\hat{X}, \hat{Y}) = (\bar{X}, \bar{Y})$, and condition (i) of Definition 5 is satisfied.

To prove condition (ii), and to be specific, let $\hat{X}_{i'}^R > \min_{k \in \mathcal{K}} \{\tilde{X}_{ki'}^R\}$, for some $i' \in \mathcal{I}^R$. Let the minimum be attained at some k' . Then

$$\hat{X}_{i'}^R > \tilde{X}_{k'i'}^R.\tag{C.5}$$

Consider inequalities (C.2) for $k = k'$. By (C.2.5), there exists a $j' \in J$ such that $\mu_{j'}^{k'} > 0$. Hence $j' \in J^*$. From (C.2.4),

$$\tilde{X}_{k'}^R \geq X_{j'}^R.\tag{C.6}$$

The inequalities (C.5) and (C.6) imply $\hat{X}_{i'}^R > X_{j'i'}^R \geq \min_{j \in J^*} \{X_{ji'}^R\} = \bar{X}_{i'}^R$, and condition (ii) of Definition 5 for the R-convex combination (\bar{X}, \bar{Y}) follows. \square

Proof of Proposition 6. Assume DMU (\hat{X}, \hat{Y}) is PR-inefficient. We need to show that there exists an optimal solution $\hat{\lambda}, \hat{\xi}$ and $\hat{\zeta}$ to program (14) such that at least one inequality in the group of constraints (14.4) and (14.5) is strict.

By Proposition 5, DMU (\hat{X}, \hat{Y}) is the R-convex combination of a finite number of observed DMUs $(X_j, Y_j), j \in J^* \subseteq J$, taken with some weights $\hat{\lambda}_j > 0, j \in J^*$, such that condition (ii) of Definition 5 is also satisfied. (In Proposition 5, the set J^* is denoted \mathcal{K} .)

Furthermore, for all $j \in J \setminus J^*$, define $\hat{\lambda}_j = 0$. Also let $\hat{\zeta} = \mathbf{0}$ and $\hat{\xi} = \mathbf{0}$. It is straightforward to verify that vectors $\hat{\lambda}, \hat{\zeta}$ and $\hat{\xi}$ are feasible in program (14). Indeed, inequalities (14.2) and (14.3) are satisfied as equalities, which follows from the fact that (\hat{X}, \hat{Y}) is the R-convex combination of observed DMUs $j \in J^*$.

To prove (14.4) and (14.5), let $j \in J^*$. (For $j \in J \setminus J^*$, $\hat{\lambda}_j = 0$, and these conditions are also true.) Then, taking into account that $\hat{\zeta}^R = \mathbf{0}$ and $\hat{\xi}^R = \mathbf{0}$,

$$\begin{aligned}Y_{jr}^R - \left[\hat{Y}_r^R + \hat{\zeta}_r^R \right] &\geq \min_{j \in J^*} \{Y_{jr}^R\} - \hat{Y}_r^R = 0, \quad \forall r \in \mathcal{O}^R, \\ X_{ji}^R - \left[\hat{X}_i^R - \hat{\xi}_i^R \right] &\leq \max_{j \in J^*} \{X_{ji}^R\} - \hat{X}_i^R = 0, \quad \forall i \in \mathcal{I}^R,\end{aligned}$$

where the right equalities in both cases follow from the definition of R-convex combination.

We have shown that $\hat{\lambda}$, $\hat{\zeta}$ and $\hat{\xi}$ is a feasible solution to program (14). Because $\hat{\zeta} = \mathbf{0}$ and $\hat{\xi} = \mathbf{0}$, the corresponding objective function of (14) is equal to zero. Therefore, $\hat{\lambda}$, $\hat{\zeta}$ and $\hat{\xi}$ is an optimal solution to (14). By condition (ii) of Definition 5, there exists a $j' \in J^*$ (for which $\hat{\lambda}_{j'} > 0$) such that either $Y_{j'r'}^R > \hat{Y}_{r'}^R$ for some $r' \in O^R$, or $X_{j'i'}^R > \hat{X}_{i'}^R$ for some $i' \in I^R$. For this j' , the corresponding inequality for output r' in the group of constraints (14.4), or the corresponding inequality for input i' in (14.5) is strict.

Conversely, assume there exists an optimal solution $\hat{\lambda}$, $\hat{\zeta}$ and $\hat{\xi}$ to program (14) such that, at least for one j , the corresponding inequality in (14.4) or (14.5) is strict. Because (\hat{X}, \hat{Y}) is strongly efficient, $\Delta^* = 0$ and, therefore, $\zeta = \mathbf{0}$ and $\xi = \mathbf{0}$. We need to prove that (\hat{X}, \hat{Y}) is PR-inefficient.

Define $J^* = \{j \in J \mid \hat{\lambda}_j > 0\}$. Let (\bar{X}, \bar{Y}) be the R-convex combination of observed DMUs (X_j, Y_j) , $j \in J^*$, taken with the weights $\hat{\lambda}_j$. Because $\zeta^V = \mathbf{0}$ and $\xi^V = \mathbf{0}$, equalities (14.2) and (14.3) imply $\bar{Y}^V = \hat{Y}^V$ and $\bar{X}^V = \hat{X}^V$. Furthermore, let $j \in J^*$. Because $\hat{\lambda}_j > 0$, inequalities (14.4) and (14.5), in which $\zeta^R = \mathbf{0}$ and $\xi^R = \mathbf{0}$, imply $Y_{jr}^R \geq \hat{Y}_r^R$, $\forall r \in O^R$, and $X_{ji}^R \leq \hat{X}_i^R$, $\forall i \in I^R$. By definition, \bar{X}^R and \bar{Y}^R satisfy (C.4). Then (\bar{X}, \bar{Y}) weakly dominates (\hat{X}, \hat{Y}) . Because, by assumption, the latter is strongly efficient, we have $(\bar{X}, \bar{Y}) = (\hat{X}, \hat{Y})$, and condition (i) of Definition 5 follows.

As assumed, at least one inequality in (14.4) and (14.5) is strict. Then there exists a $j' \in J^*$ and an $r' \in O^R$ such that $Y_{j'r'}^R > \hat{Y}_{r'}^R$, or there exists a $j' \in J^*$ and an $i' \in I^R$ such that $X_{j'i'}^R < \hat{X}_{i'}^R$. In both cases condition (ii) of Definition 5 follows. \square

Proof of Proposition 7. Assume DMU (\hat{X}, \hat{Y}) is PR-inefficient. We need to show that $\Delta_1 > 0$. By Proposition 5, DMU (\hat{X}, \hat{Y}) is the R-convex combination of a finite number of observed DMUs (X_j, Y_j) , $j \in J^* \subseteq J$, taken with some weights $\hat{\lambda}_j > 0$, $j \in J^*$, which satisfies condition (ii) of Definition 5. (In Proposition 5, the set J^* is denoted \mathcal{K} .)

Define $\hat{\lambda}_j = 0$, for all $j \in J \setminus J^*$, and substitute vector $\hat{\lambda}$ in (16). Because (\hat{X}, \hat{Y}) is the R-convex combination of observed DMUs (X_j, Y_j) , $j \in J^*$, equations (16.2) and (16.3) are true. To define vectors $\hat{\zeta}_j^R$ and $\hat{\xi}_j^R$, consider two cases. If $j \in J \setminus J^*$, we let $\hat{\zeta}_j^R = \mathbf{0}$ and $\hat{\xi}_j^R = \mathbf{0}$, as required by (16.6). Let $j \in J^*$. Then $\hat{\lambda}_j > 0$ and, to satisfy equations (16.4) and (16.5), we define vectors $\hat{\zeta}_j$ and $\hat{\xi}_j$ as follows:

$$\begin{aligned} \hat{\zeta}_{jr}^R &= Y_{jr}^R - \hat{Y}_r^R \geq 0, & \forall r \in O^R, \\ \hat{\xi}_{ji}^R &= \hat{X}_i^R - X_{ji}^R \leq 0, & \forall i \in I^R. \end{aligned} \tag{C.7}$$

Vectors $\hat{\lambda}$, $\hat{\zeta}^R$ and $\hat{\xi}^R$ represent a feasible solution to program (16). By condition (ii) of Definition 5, for at least one $r \in O^R$ or $i \in I^R$, at least one of the inequalities in (C.7) is strict. Then the objective function in (16) is strictly positive, and the optimal $\Delta_1 > 0$.

Conversely, let $\hat{\lambda}$, $\hat{\zeta}_j^R$, $\hat{\xi}_j^R$, $j \in J$, be an optimal solution to program (16), and let $\Delta_1 > 0$. We need to show that (\hat{X}, \hat{Y}) is PR-inefficient. Define $J^* = \{j \mid \hat{\lambda}_j > 0, j \in J\}$. Let (\bar{X}, \bar{Y}) be the R-convex combination of observed DMUs (X_j, Y_j) , $j \in J^*$, taken with the weights $\hat{\lambda}_j$. Then equalities (16.2) and (16.3) imply $\bar{Y}^V = \hat{Y}^V$ and $\bar{X}^V = \hat{X}^V$. Because $\hat{\lambda}_j > 0$, $\forall j \in J^*$, (16.4) and (16.5) imply that \bar{X}^R and \bar{Y}^R satisfy (C.4). Therefore, (\bar{X}, \bar{Y}) weakly dominates (\hat{X}, \hat{Y}) . By assumption, the latter is strongly efficient. Therefore, we have $(\bar{X}, \bar{Y}) = (\hat{X}, \hat{Y})$, and condition (i) of Definition 5 follows.

Because $\Delta_1 > 0$, there exist a $j' \in J^*$ and $r' \in O^R$ such that $\hat{\zeta}_{j'r'}^R > 0$, or there exist a $j' \in J^*$ and an $i' \in I^R$ such that $\hat{\xi}_{j'i'}^R > 0$. In the former case, (16.4) implies that $Y_{j'r'}^R > \hat{Y}_{r'}^R$. In the latter case, by (16.5), $X_{j'i'}^R < \hat{X}_{i'}^R$, and condition (ii) of Definition 5 follows. \square

Proof of Proposition 8. Assume DMU (\hat{X}, \hat{Y}) is PR-inefficient. We need to show that $\Delta_2 > 0$. By Proposition 5, DMU (\hat{X}, \hat{Y}) is the R-convex combination of a finite number of observed DMUs (X_j, Y_j) , $j \in J^* \subseteq J$, taken with some weights $\hat{\lambda}_j > 0$, $j \in J^*$, which satisfies condition (ii) of Definition 5. (In Proposition 5, the set J^* is denoted \mathcal{K} .) Clearly, $J^* \subseteq J_0$.

Define $\hat{\lambda}_j = 0$, for all $j \in J \setminus J^*$, and substitute vector $\hat{\lambda}$ in (17). Because (\hat{X}, \hat{Y}) is the R-convex combination of observed DMUs (X_j, Y_j) , $j \in J^*$, equations (17.2) and (17.3) are true. Furthermore, define

$$\begin{aligned}\hat{\zeta}_r^R &\triangleq \sum_{j \in J_0} \hat{\lambda}_j Y_{jr}^R - \hat{Y}_r^R \geq \min_{j \in J_0} \{Y_{jr}^R\} - \hat{Y}_r^R = 0, & \forall r \in O^R, \\ \hat{\xi}_i^R &\triangleq \hat{X}_i^R - \sum_{j \in J_0} \hat{\lambda}_j X_{ji}^R \leq \hat{X}_i^R - \max_{j \in J_0} \{X_{ji}^R\} = 0, & \forall i \in I^R,\end{aligned}\tag{C.8}$$

where the right equalities in the two cases follow from the definition of R-convex combination.

Therefore, vectors $\hat{\lambda}$, $\hat{\zeta}^R$ and $\hat{\xi}^R$ represent a feasible solution to program (17). By condition (ii) of Definition 5, for at least one $r \in O^R$ or $i \in I^R$, the inequality in (C.8) is strict. Then the objective function in (17) is strictly positive, and the optimal $\Delta_2 > 0$.

Conversely, let $\hat{\lambda}$, $\hat{\zeta}^R$, $\hat{\xi}^R$ be an optimal solution to program (17), and let $\Delta_2 > 0$. We need to prove that (\hat{X}, \hat{Y}) is PR-inefficient.

Define $J^* = \{j \in J_0 \mid \hat{\lambda}_j > 0\}$. Let (\bar{X}, \bar{Y}) be the R-convex combination of observed DMUs (X_j, Y_j) , $j \in J^*$, taken with the weights $\hat{\lambda}_j$.

First note that equalities (17.2) and (17.3) imply $\bar{Y}^V = \hat{Y}^V$ and $\bar{X}^V = \hat{X}^V$. By definition of the set J_0 , \bar{X}^R and \bar{Y}^R satisfy (C.4). Then (\bar{X}, \bar{Y}) weakly dominates (\hat{X}, \hat{Y}) . Because, by assumption, the latter is strongly efficient, we have $(\bar{X}, \bar{Y}) = (\hat{X}, \hat{Y})$, and condition (i) of Definition 5 follows.

Because $\Delta_2 > 0$, there exists an $r' \in O^R$ such that $\hat{\zeta}_{r'}^R > 0$, or there exists an $i' \in I^R$ such that $\hat{\xi}_{i'}^R > 0$. In the former case, (17.4) implies that there exists a $j' \in J^*$ such that $Y_{j'r'}^R > \hat{Y}_{r'}^R$. In the latter case, by (17.5), there exists a $j' \in J^*$ such that $X_{j'i'}^R < \hat{X}_{i'}^R$, and condition (ii) of Definition 5 follows. \square

Proof of Proposition 10. The proof follows closely the proof of Proposition 8. Assume DMU (\hat{X}, \hat{Y}) is PR-inefficient. We need to show that $\Delta_3 > 0$.

By Proposition 9, DMU (\hat{X}, \hat{Y}) is the R-conical combination of a finite number of observed DMUs (X_j, Y_j) , $j \in J^* \subseteq J$, taken with some weights $\hat{\lambda}_j > 0$, $j \in J^*$, which satisfies condition (ii) of Definition 5. Note that, in Proposition 5, the set J^* is denoted \mathcal{K} . Clearly, $J^* \subseteq J_0$.

Define $\hat{\lambda}_j = 0$, for all $j \in J \setminus J^*$. Denote

$$\Lambda_0 = \sum_{j \in J_0} \hat{\lambda}_j > 0.\tag{C.9}$$

Substitute vector $\hat{\lambda}$ in (17). Because (\hat{X}, \hat{Y}) is the R-conical combination of observed DMUs (X_j, Y_j) , $j \in J^*$, equations (18.2) and (18.3) are true. Furthermore, define

$$\begin{aligned}\hat{\zeta}_r^R &\triangleq \sum_{j \in J_0} \hat{\lambda}_j (Y_{jr}^R - \hat{Y}_r^R) \geq \Lambda_0 \left(\min_{j \in J_0} \{Y_{jr}^R\} - \hat{Y}_r^R \right) = 0, & \forall r \in O^R, \\ \hat{\xi}_i^R &\triangleq \sum_{j \in J_0} \hat{\lambda}_j (\hat{X}_i^R - X_{ji}^R) \leq \Lambda_0 \left(\hat{X}_i^R - \max_{j \in J_0} \{X_{ji}^R\} \right) = 0, & \forall i \in I^R,\end{aligned}\tag{C.10}$$

where the right equalities in the two cases follow from Definition 4 of R-conical combination and, in particular, from conditions (1.3) and (1.4).

By construction, vectors $\hat{\lambda}$, $\hat{\zeta}^R$ and $\hat{\xi}^R$ represent a feasible solution to program (18). By condition (ii) of Definition 5, for at least one $r \in O^R$ or $i \in I^R$, the inequality in (C.10) is strict. Taking into account (C.9), the corresponding slack $\hat{\zeta}_r^R$ or $\hat{\xi}_i^R$ defined by (C.10) is positive. Therefore, the objective function in (18) is strictly positive, and the optimal $\Delta_3 > 0$.

Conversely, let $\hat{\lambda}$, $\hat{\zeta}^R$, $\hat{\xi}^R$ be an optimal solution to program (18), and let $\Delta_3 > 0$. We need to show that (\hat{X}, \hat{Y}) is PR-inefficient.

Define $J^* = \{j \in J_0 \mid \hat{\lambda}_j > 0\}$. By the assumption that $\hat{Y}^V \neq \mathbf{0}$ (see footnote 20), constraints (18.2) imply (C.9), and the set $J^* \neq \emptyset$. Let (\bar{X}, \bar{Y}) be the R-conical combination of observed DMUs (X_j, Y_j) , $j \in J^*$, taken with the weights $\hat{\lambda}_j$.

First note that equalities (18.2) and (18.3) imply $\bar{Y}^V = \hat{Y}^V$ and $\bar{X}^V = \hat{X}^V$. Also, by definition of the set J_0 , \bar{X}^R and \bar{Y}^R satisfy (C.4). Then (\bar{X}, \bar{Y}) weakly dominates (\hat{X}, \hat{Y}) . Because, by assumption, the latter is strongly efficient, we have $(\bar{X}, \bar{Y}) = (\hat{X}, \hat{Y})$, i.e., DMU (\hat{X}, \hat{Y}) is a conical combination (\bar{X}, \bar{Y}) of observed DMUs (X_j, Y_j) , $j \in J^*$, and condition (i) of Definition 5 follows.

Furthermore, because $\Delta_3 > 0$, there exists an $r' \in O^R$ such that $\hat{\zeta}_{r'}^R > 0$, or there exists an $i' \in I^R$ such that $\hat{\xi}_{i'}^R > 0$. In the former case, (18.4) implies that there exists a $j' \in J^*$ such that $Y_{j'r'}^R > \hat{Y}_{r'}^R$. In the latter case, by (18.5), there exists a $j' \in J^*$ such that $X_{j'i'}^R < \hat{X}_{i'}^R$, and condition (ii) of Definition 5 follows.

We have shown that DMU (\hat{X}, \hat{Y}) satisfies the conditions of Proposition 9 and is, therefore, PR-inefficient. \square

References

- Ali, A.I., Seiford, L.M. (1993). The mathematical programming approach to efficiency analysis. In H.O. Fried, C.A.K. Lovell, S.S. Schmidt (Eds.), *The measurement of productive efficiency: Techniques and applications* (pp. 120-159). New York: Oxford University Press.
- Banker, R.D., Charnes, A., Cooper, W.W. (1984). Some models for estimating technical and scale efficiencies 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.
- Blackburn, V., Brennan, S., Ruggiero, J. (2014). Measuring efficiency in Australian schools: A preliminary analysis. *Socio-Economic Planning Sciences*, 48(1), 4–9.
- Chambers, R.G., Chung, Y., Färe, R. (1996). Benefit and distance functions. *Journal of Economic Theory*, 70(2), 407–419.
- Chambers, R.G., Chung, Y., Färe, R. (1998). Profit, directional distance functions, and Nerlovian efficiency. *Journal of Optimization Theory and Applications*, 98(2), 351–364.
- 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., Seiford, L.M., Tone, K. (2007). *Data envelopment analysis. A comprehensive text with models, applications, references and DEA-Solver software* (2nd ed.). New York: Springer Science + Business Media.
- 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). Amsterdam: North-Holland.
- Emrouznejad, A., Amin, G.R. (2009) DEA models for ratio data: Convexity consideration. *Applied Mathematical Modelling*, 33(1), 486–498.
- Haelermans, C., Ruggiero, J. (2013). Estimating technical and allocative efficiency in the public sector: A nonparametric analysis of Dutch schools. *European Journal of Operational Research*, 227(1), 174–181.
- Johnson, A., Ruggiero, J. (2014). Nonparametric measurement of productivity and efficiency in education. *Annals of Operations Research*, 221(1), 197–210.
- Olesen, O.B., Petersen, N.C., Podinovski, V.V. (2015). Efficiency analysis with ratio measures. *European Journal of Operational Research*, 245(2), 446–462.
- Podinovski, V.V. (2005). Selective convexity in DEA models. *European Journal of Operational Research*, 161(2), 552–563.
- Ruggiero, J. (1996). On the measurement of technical efficiency in the public sector. *European Journal of Operational Research*, 90(3), 553–565.