European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

Production trade-offs in models of data envelopment analysis with ratio inputs and outputs: An application to schools in England

Victor V. Podinovski[a,*], Junlin Wu[b], Nikolaos Argyris[a]

[a] Loughborough Business School, Loughborough University, Leicestershire LE11 3TU, UK
[b] School of Business, Operations and Strategy, University of Greenwich, London SE10 9LS, UK

ARTICLE INFO

Article history:
Received 12 February 2023
Accepted 11 August 2023
Available online xxx

Keywords:
Data envelopment analysis
Production technology
Ratio inputs and outputs
Production trade-offs
Secondary schools

ABSTRACT

In applications of data envelopment analysis (DEA), the inputs and outputs representing environmental and quality characteristics of the production process are often stated in the form of percentages, ratios and averages, collectively referred to as ratio measures. It is known that the conventional variable and constant returns-to-scale (VRS and CRS) DEA models cannot correctly incorporate such ratio inputs and outputs. This problem has been addressed by the development of Ratio-VRS and Ratio-CRS (R-VRS and R-CRS) models suitable for the incorporation of both volume and ratio inputs and outputs. Such models may, however, depending on the application, lack sufficient discriminatory power. In this paper we address this issue by developing a further extension of the R-VRS and R-CRS models (the latter with the most common fixed type of ratio inputs and outputs) by allowing the specification of production trade-offs between volume inputs and outputs, and, similarly, between ratio measures. As in the case of conventional VRS and CRS models in which the role of production trade-offs is well understood, the specification of such trade-offs in the R-VRS and R-CRS production technologies leads to their controlled expansion and results in improved efficiency discrimination of the resulting DEA models. We illustrate the application of the proposed methodology by the assessment of efficiency of a large sample of secondary schools in England.

1. Introduction

Applications of data envelopment analysis (DEA) often incorporate inputs and outputs stated in the form of ratios. Such *ratio measures* may represent various percentages and ratios obtained by division of volume measures whose values are often unknown to the analyst. For example, in the context of school education, which we consider in the application, ratio measures may represent socio-economic characteristics of the pupil intake, such as percentage of school pupils eligible for free school meals. Ratio measures may also represent percentage of pupils achieving good results in exams and those proceeding to higher education after the graduation. More broadly, ratio measures are often used as environmental and quality characteristics of the production process.

* Corresponding author.
E-mail addresses: v.podinovski@lboro.ac.uk (V.V. Podinovski), j.wu@greenwich.ac.uk (J. Wu), n.argyris@lboro.ac.uk (N. Argyris).

1.1. Existing approaches

It is now well known that ratio inputs and outputs cannot be correctly incorporated in the conventional variable and constant returns-to-scale (VRS and CRS) DEA models of Charnes, Cooper, & Rhodes (1978) and Banker, Charnes, & Cooper (1984). The earlier concerns for the applicability of ratios in DEA were raised, for example, by Dyson et al. (2001) and Cooper, Seiford and Tone (2007). The examples given by Emrouznejad & Amin (2009) and Olesen, Petersen, & Podinovski (2015) show that the incorporation of ratio inputs and outputs in the standard VRS and CRS models is inconsistent with the assumption that the underlying technology is convex. In the case of CRS, ratio measures are usually also inconsistent with the assumption of scalability of data (Olesen et al., 2015).

The Ratio-VRS and Ratio-CRS (R-VRS and R-CRS) models of technology developed by Olesen et al. (2015) address the described problem and allow the incorporation of both volume and ratio inputs and outputs as native types of data. In this approach, the standard assumption that the production technology is convex is

https://doi.org/10.1016/j.ejor.2023.08.019

replaced by the assumption of selective convexity of Podinovski (2005). The R-VRS and R-CRS models allow convex combinations of decision making units (DMUs) in which the volume measures are combined in the conventional way. At the same time, the ratio measures are taken at the most conservative level that may be obtained by accounting for all possible (but typically unknown) values of their numerators and denominators—see Papaioannou & Podinovski (2023) for additional results concerning the impact of different types of information about the numerators and denominators on the model of technology.

The R-CRS model of Olesen et al. (2015) additionally allows proportional scaling of the volume inputs and outputs. The treatment of ratio measures is more nuanced and depends on their assumed type. From a practical perspective, the most common is the *fixed* type of ratio inputs and outputs. These represent environmental and quality characteristics of the production process that can be assumed constant while the volume inputs and outputs characterizing the quantity of resources and production levels are scaled up and down. Olesen et al. (2015) refer to such R-CRS technologies with fixed ratio inputs and outputs as the F-CRS technologies.

1.2. Motivation

While the R-VRS and F-CRS models allow for appropriately capturing ratio inputs and outputs in the specification of the production technology, their use in practice can be hindered by the issue of lack of discriminatory power. This is a well-documented finding for traditional VRS and CRS DEA models, in response to which different rules of thumb for the number of inputs and outputs required for acceptable discrimination of the model exist (Dyson et al., 2001; Cooper et al. 2007). Since the R-VRS technology is a subset of the conventional VRS technology based on the same input and output data (including ratio data), the efficiency scores can only increase when moving from the latter to the former. Therefore, the discriminating power of the R-VRS model is generally worse than that of the VRS model. Although the F-CRS technology is generally not a subset of the CRS technology,[1] computations show that it also generally lacks in discriminating power compared to the standard CRS model.

A well-established methodology to mitigate against this problem in the *conventional* VRS and CRS setting involves incorporating value or expert judgments in the specification of production technology, via means of specifying production trade-offs. These were originally developed by Podinovski (2004) as the dual forms of weight restrictions in the standard multiplier VRS and CRS models (see, e.g., Allen, Athanassopoulos, Dyson, & Thanassoulis, 1997, and Thanassoulis, Portela, & Despić, 2008). Such production trade-offs are interpretable as simultaneous changes to the inputs or outputs that are assumed technologically possible for any DMU in the technology. An example of such a judgement, aligned with the application to schools considered in our paper, is the statement that 1 extra teacher and £20,000 of non-pay school expenses is a sufficient compensation for a school to accept at least 10 extra pupils. Such production trade-offs result in a controlled and meaningful expansion of the production technology, leading to potentially lower efficiency scores and therefore improved discriminatory power.

A question therefore arises as to whether it is possible to expand the R-VRS and F-CRS models and improve their discriminating power by defining similar production trade-offs involving both volume and ratio inputs and outputs. In our paper, we give a positive answer to this question and illustrate the results by an application.

1.3. Contribution

In this paper, we develop extensions of the original R-VRS and F-CRS technologies of Olesen et al. (2015) by the incorporation of production trade-offs involving volume and ratio inputs and outputs. Any new DMU in such extended technologies can be explained as a modification of some DMU in the original R-VRS and F-CRS technology by the application of the assumed trade-offs.

It should be noted that the development of such models is not a straightforward modification of the R-VRS and F-CRS technologies, for two reasons.

First, the R-VRS and F-CRS technologies are not convex, the envelopment models based on them are not linear and their standard dual multiplier models are undefined. This means that the notion of production trade-offs does not arise as the dual form of weight restrictions and needs to be developed as a stand-alone concept, based on the envelopment form only. We achieve this by providing a full axiomatic development of the R-VRS and F-CRS technologies with production trade-offs entirely in the primal (envelopment) space.

Second, we distinguish between the production trade-offs stated for volume inputs and outputs (as illustrated above), and trade-offs stated for ratio measures. An example of the latter is the judgement that, if the percentage of pupils achieving good results on entry to school increases by 1%, any school should be able to improve the percentage of pupils achieving good results on exit by at least 0.5%. It turns out that the production trade-offs stated for the ratio measures require a different modelling approach compared to the trade-offs stated for the volume measures.

As in the case of conventional VRS and CRS models, the incorporation of production trade-offs in the R-VRS and F-CRS technologies results to their expansion. This in turn leads to improved discrimination on efficiency of the resulting expanded models. We illustrate the usefulness of the new R-VRS and F-CRS models with production trade-offs by an application to a large sample of secondary schools in England. Computational results confirm that the specification of production trade-offs leads to a noticeable improvement of the discrimination on efficiency.

We proceed as follows. In Section 2, we briefly outline the idea of the R-VRS technology of Olesen et al. (2015). In Section 3, we introduce the notion of production trade-offs specified either for the volume or ratio inputs and outputs. In Sections 4 and 5, we use an axiomatic approach to develop the R-VRS and F-CRS technologies with production trade-offs. In Section 6, we discuss DEA models based on the new technologies and approaches to their solution. In Section 7, we consider an application to secondary schools in England. Concluding remarks are given in Section 8. The proofs of all mathematical statements are given in Appendix A.

2. Preliminaries

In this section, we provide a brief introduction to the R-VRS technology developed by Olesen et al. (2015) and its axiomatic foundations.

2.1. Notation

Let $T \in \mathbb{R}_+^{m+s}$ be a production technology with the set $\mathrm{I} = \{1, \ldots, m\}$ of inputs and the set $\mathrm{O} = \{1, \ldots, s\}$ of outputs. Both sets may generally include volume and ratio measures that require different treatment.

Denote I^V and O^V the subsets of volume inputs and outputs, respectively. Similarly, denote I^R and O^R the subsets of ratio inputs

[1] To see this, note that the F-CRS technology allows selective proportional increase of volume measures only, which does not require increasing fixed ratio inputs. The resulting DMUs are generally outside the standard CRS technology which only allows simultaneous proportional increase of all inputs and outputs.

and outputs. We obviously have $I = I^V \cup I^R$ and $O = O^V \cup O^R$, and $I^V \cap I^R = \varnothing$ and $O^V \cap O^R = \varnothing$. We assume that there is at least one input and at least one output, i.e., the sets I and O are not empty, although some of their subsets of volume and ratio measures may be empty.

Denote $|I^V| = m^V$ and $|I^R| = m^R$ the number of volume and ratio inputs, respectively. Then the overall number of inputs is $m = m^V + m^R$. Similarly, we denote $|O^V| = s^V$ and $|O^R| = s^R$ the number of volume and ratio outputs, respectively. The overall number of outputs is $s = s^V + s^R$.

Decision making units (DMUs) are elements of technology T. They may be stated in the form (X, Y), where $X \in \mathbb{R}^m_+$ and $Y \in \mathbb{R}^s_+$ are the vectors of inputs and outputs, or in the more detailed form that reflects the division of the inputs and outputs into the volume and ratio measures, as follows:

$$(X, Y) = \left(X^V, X^R, Y^V, Y^R\right), \tag{1}$$

where $X^V \in \mathbb{R}^{m^V}_+$, $X^R \in \mathbb{R}^{m^R}_+$, $Y^V \in \mathbb{R}^{s^V}_+$ and $Y^R \in \mathbb{R}^{s^R}_+$.

Suppose that we have n observed DMUs. Introducing the index set $J = \{1, \ldots, n\}$, we state the observed DMUs as $(X_j, Y_j) = (X^V_j, X^R_j, Y^V_j, Y^R_j)$, $j \in J$.

In many applications, ratio measures often (but not always) have natural upper bounds, typically either unity or 100%. Let $\bar{X}^R$ and $\bar{Y}^R$ be the vectors of upper bounds on the ratio inputs and outputs. If an upper bound on an input $i \in I^R$ or output $r \in O^R$ is not specified, we formally take the corresponding upper bound $\bar{X}^R_i$ or $\bar{Y}^R_r$ equal to $+\infty$. All DMUs (1) in technology T are naturally assumed to satisfy the two vector inequalities:

$$X^R \leq \bar{X}^R \quad \text{and} \quad Y^R \leq \bar{Y}^R. \tag{2}$$

(In this paper, including in inequalities (2), vector inequalities mean that the specified inequality is true for each component. For example, the vector inequality $X^R \leq \bar{X}^R$ means that the scalar inequality $X^R_i \leq \bar{X}^R_i$ is true for all $i \in I^R$.)

2.2. Basic axioms

Banker et al. (1984) show that the conventional VRS technology is the smallest technology (in the sense of the minimum extrapolation principle) that is generated by the set of observed DMUs and satisfies the axioms of free disposability and convexity, as specified formally below. It is known that these two axioms cannot be assumed if some inputs or outputs are stated as ratio measures (Emrouznejad & Amin, 2009; Olesen et al., 2015).

To provide an axiomatic foundation of the R-VRS technology (and the R-CRS technology in a further development), Olesen et al. (2015) modify the axioms of Banker et al. (1984). The axiom of free disposability requires a simple modification stating that the worsening of the inputs and outputs of any DMU in technology T is possible as long as the resulting DMU remains within the bounds (2) on the ratio measures.

Furthermore, Olesen et al. (2015) note that, although taking convex combinations of DMUs with ratio inputs and outputs would be incorrect, there is a special case in which such convex combinations are justified. Let us illustrate this by a simple example.

Example 1. Suppose that we want to define a convex combination of two schools A and B taken, to be specific, with equal weights 0.5. We define all volume inputs and outputs of the combined school C as the simple average of the corresponding volume inputs and outputs of schools A and B. However, we cannot treat ratio measures in the same way. Indeed, let the percentage of pupils achieving good results on exit at schools A and B be equal to p_A and p_B. Then, for the combined school C, the percentage p_C of such pupils may be anywhere in the range between p_A and p_B, which depends on the (unknown to us) numerators and denominators that define such percentages. However, if $p_A = p_B$, we always have $p_C = p_A = p_B$.

This example shows that we may take convex combinations of DMUs, provided they have the same subvectors of ratio inputs and outputs. This observation motivates (Olesen et al., 2015) to replace the standard axiom of convexity used by Banker et al. (1984) by the axiom of selective convexity introduced by Podinovski (2005). This axiom allows convexity with respect to the selected sets of inputs and outputs (sets I^V and O^V in our case), assuming the remaining inputs and outputs (in the sets I^R and O^R) are identical for the combined DMUs.

Olesen et al. (2015) state the following three axioms as the foundation of the R-VRS technology.

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

Axiom 2 Free disposability**.** Let $(X, Y) = (X^V, X^R, Y^V, Y^R) \in T$ and let $(\tilde{X}, \tilde{Y}) = (\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$. Suppose that the subvectors $\tilde{X}^R$ and $\tilde{Y}^R$ are within the bounds stated by inequalities (2) and that $\tilde{Y} \leq Y$, $\tilde{X} \geq X$. Then $(\tilde{X}, \tilde{Y}) \in T$.

Axiom 3 Selective convexity**.** Let $(\tilde{X}, \tilde{Y}) \in T$ and $(\hat{X}, \hat{Y}) \in T$, and let

$$\tilde{X}^R = \hat{X}^R \quad \text{and} \quad \tilde{Y}^R = \hat{Y}^R. \tag{3}$$

Then $\gamma(\tilde{X}, \tilde{Y}) + (1 - \gamma)(\hat{X}, \hat{Y}) \in T$, for any $\gamma \in [0, 1]$.

2.3. The R-VRS technology

Following the minimum extrapolation principle used by Banker et al. (1984), Olesen et al. (2015) define the R-VRS technology $T^{\mathrm{R}}_{\mathrm{VRS}}$ as the intersection of all technologies that satisfy Axioms 1–3, i.e., as the smallest technology that satisfies these axioms.

As proved by Olesen et al. (2015), technology $T^{\mathrm{R}}_{\mathrm{VRS}}$ coincides with the set of all DMUs $(X, Y) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ for which there exists a vector $\lambda \in \mathbb{R}^n$ such that the following conditions are true:

$$\sum_{j=1}^{n} \lambda_j Y^V_j \geq Y^V, \tag{4a}$$

$$\sum_{j=1}^{n} \lambda_j X^V_j \leq X^V, \tag{4b}$$

$$\lambda_j \left(Y^R_j - Y^R\right) \geq \mathbf{0}, \qquad \forall j \in J, \tag{4c}$$

$$\lambda_j \left(X^R_j - X^R\right) \leq \mathbf{0}, \qquad \forall j \in J, \tag{4d}$$

$$\sum_{j=1}^{n} \lambda_j = 1, \tag{4e}$$

$$X^R \leq \bar{X}^R, \tag{4f}$$

$$Y^R \leq \bar{Y}^R, \tag{4g}$$

$$\lambda \geq \mathbf{0}. \tag{4h}$$

(In the statement (4) and elsewhere in this paper, we use bold notation $\mathbf{0}$ for vectors of zeros whose dimensions are clear from the context.)

The role of conditions (4c) and (4d) is straightforward. Namely, for every $j \in J$, if $\lambda_j > 0$, then both vector inequalities $Y^R_j \geq Y^R$ and $X^R_j \leq X^R$ must be satisfied. Let, for example, the ratio inputs represent the environment in which the DMUs operate and let the ratio outputs describe the quality of the outputs produced by the DMUs. Then conditions (4c) and (4d) mean that the convex combinations of the volume inputs and outputs, taken with the weights λ_j on the left-hand side of inequalities (4a) and (4b), include only those observed DMUs (X_j, Y_j), $j \in J$, that do not operate in a more favourable environment than DMU (X, Y) and whose quality of production is at least as good as that of DMU (X, Y).

3. Production trade-offs

Podinovski (2004) defines production trade-offs as value judgements stating that certain simultaneous changes to the inputs and outputs are technologically feasible throughout the entire technology. In the case of conventional VRS and CRS technologies, production trade-offs are dual to weight restrictions in the multiplier models based on these technologies.

Our objective is to extend the notion of production trade-offs to the R-VRS technology. Two differences with the case of conventional VRS technology are worth highlighting. First, the R-VRS technology is generally not convex and the input and output-oriented envelopment programs based on it are not linear programs. Consequently, such programs do not have dual multiplier forms in which weight restrictions could be incorporated. As a result, we introduce production trade-offs for the R-VRS technology without reference to weight restrictions. Second, it turns out that the mechanism by which the trade-offs are specified for volume inputs and outputs is different from the approach required for the trade-offs involving ratio measures.

Following notation of Podinovski (2004), a production trade-off between volume inputs and outputs can be stated by the pair of vectors $P^V \in \mathbb{R}^{m^V}$ and $Q^V \in \mathbb{R}^{s^V}$, which describe simultaneous changes to the vectors of volume inputs and outputs of the DMUs in the technology. Similarly, a production trade-off between ratio inputs and outputs can be stated by the pair of vectors $P^R \in \mathbb{R}^{m^R}$ and $Q^R \in \mathbb{R}^{s^R}$, which describe simultaneous changes to the vectors of ratio inputs and outputs of the DMUs. Components of these vectors can be positive, negative or equal to zero.

Suppose we have specified $K \geq 0$ production trade-offs between volume inputs and outputs. We state these as follows:

$$\left(P_t^V, Q_t^V\right), \quad t = 1, \ldots, K. \tag{5}$$

Similarly, we state $L \geq 0$ production trade-offs between ratio inputs and outputs as

$$\left(P_l^R, Q_l^R\right), \quad l = 1, \ldots, L. \tag{6}$$

Example 2. To illustrate the idea of production trade-offs, we consider an example in the context of secondary schools, which is aligned with the application considered in Section 7. Let the vector of inputs be $X = (x_1^V, x_2^V, x_3^R)^\top$. The volume inputs x_1^V and x_2^V represent the number of teachers and school expenses, respectively. The ratio input x_3^R is the percentage of pupils with good academic results on entry to the school. Similarly, let $Y = (y_1^V, y_2^R)^\top$ be the vector of outputs, where the volume output y_1^V represents the number of pupils and the ratio output y_2^R is the percentage of pupils achieving good results on exit from school.

In the described setting, we may consider several value judgements and state them as production trade-offs. (Note that the exact values will obviously need to be assessed and justified in any particular application. The simple examples provided below are intended for conceptual purposes only, as an illustration of the idea of a trade-off.)

First, we may judge that it is technologically possible for any school to increase the intake by extra 10 pupils, provided the school recruits an additional teacher and is given an extra budget of £20,000 per year. The corresponding trade-off is now stated as $P_1^V = (1, £20{,}000)^\top$ and $Q_1^V = (10)$.

Second, we may assume that schools can compensate a loss of one teacher by a bought-in teacher using their budgets, and specify a trade-off for this. Suppose that £50,000 should be sufficient to pay for a substitute teacher. We can now state the second volume trade-off as $P_2^V = (-1, £50{,}000)^\top$ and $Q_2^V = (0)$. Note that the zero component of the vector Q_2^V means that there is no change to the number of pupils.

Third, we may make a conservative judgement that, if the percentage of pupils achieving good results on entry to school increases by 1%, the percentage of pupils achieving good results on exit should increase by at least 0.5%. On the other hand, if the former percentage is reduced by 1%, this should not result in more than 1% fewer pupils achieving good results on exit. These two judgements are stated in the form of production trade-offs as follows: $P_1^R = (1\%)$, $Q_1^R = (0.5\%)$, and $P_2^R = (-1\%)$, $Q_2^R = (-1\%)$, respectively.

In line with Podinovski (2004), we consider production trade-offs as conservative judgements that can be applied to any DMU in the technology and any (not necessarily integer) number of times, as long as the inputs and outputs of the resulting DMU remain nonnegative and within the bounds (2) specified for the ratio measures. For example, applying trade-off $P_1^R = (1\%)$ and $Q_1^R = (0.5\%)$ from Example 2 three times, we conclude that it is technologically possible for any school to increase the percentage of pupils achieving good results on exit by $3 \times 0.5\% = 1.5\%$, if the percentage of pupils with good results on entry increases by $3 \times 1\% = 3\%$.

Let us provide a formal statement of the assumption that the production trade-offs represent technologically feasible simultaneous changes to the vectors of inputs and outputs. Let multipliers $\pi_t \geq 0$, $t = 1, \ldots, K$, and $\rho_l \geq 0$, $l = 1, \ldots, L$, represent the proportions in which we apply the volume and ratio trade-offs (5) and (6), respectively. The following axiom is a generalization of a similar axiom stated by Podinovski (2004) for the conventional VRS and CRS technologies.

Axiom 4 Feasibility of production trade-offs**.** Let $(X^V, X^R, Y^V, Y^R) \in T$. Consider any scalars $\pi_t \geq 0$, for all $t = 1, \ldots, K$, and $\rho_l \geq 0$, for all $l = 1, \ldots, L$. Define DMU $(\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R)$, where

$$\begin{aligned} \tilde{Y}^V &= Y^V + \sum_{t=1}^{K} \pi_t Q_t^V, \\ \tilde{X}^V &= X^V + \sum_{t=1}^{K} \pi_t P_t^V, \\ \tilde{Y}^R &= Y^R + \sum_{l=1}^{L} \rho_l Q_l^R, \\ \tilde{X}^R &= X^R + \sum_{l=1}^{L} \rho_l P_l^R. \end{aligned} \tag{7}$$

Further assume that $\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R \geq \mathbf{0}$ and that the subvectors $\tilde{X}^R$ and $\tilde{Y}^R$ are within the bounds (2). Then $\left(\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R\right) \in T$.

Remark 1. We use production trade-offs (5) and (6) to describe technologically feasible changes to either volume or ratio inputs and outputs, but not simultaneous changes to both types. There are two reasons for this.

First, using Example 2 for illustration, we are not stating that increasing the number of teachers (input x_1^V) by one may lead to a certain improvement of the percentage of pupils graduating from school with good results (output y_2^R). This would appear problematic because one additional teacher may make a big difference for output y_2^R in a small school and a small difference in a large school.

Second, as shown below, the volume and ratio trade-offs (5) and (6) are incorporated differently in the R-VRS technology. It is not clear how a mixed trade-off involving changes to both volume and ratio measures could be incorporated. Because, as noted, the practical meaning of such trade-offs is questionable, this possibility is not considered.

4. The R-VRS technology with production trade-offs

In this section, we obtain an extension of the R-VRS technology of Olesen et al. (2015) by the incorporation of production trade-offs (5) and (6). We denote this technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ and formally derive it from the stated Axioms 1–4.

Definition 1. Technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ is the intersection of all sets $T \subset \mathbb{R}^m_+ \times \mathbb{R}^s_+$ that satisfy Axioms 1–4.

It is straightforward to prove that technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ satisfies all Axioms 1–4. It can, therefore, be regarded as the smallest technology that satisfies these axioms. The following theorem provides an equivalent explicit statement of this technology.

Theorem 1. *Technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ is the set of all DMUs $(X,Y) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ for which there exist vectors $\lambda \in \mathbb{R}^n$, $\pi \in \mathbb{R}^K$ and scalars ρ_{jl}, $j \in J$, $l = 1,\dots,L$, such that*

$$\sum_{j=1}^{n} \lambda_j Y_j^V + \sum_{t=1}^{K} \pi_t Q_t^V \geq Y^V, \tag{8a}$$

$$\sum_{j=1}^{n} \lambda_j X_j^V + \sum_{t=1}^{K} \pi_t P_t^V \leq X^V, \tag{8b}$$

$$\lambda_j \left(\left[Y_j^R + \sum_{l=1}^{L} \rho_{jl} Q_l^R \right] - Y^R \right) \geq \mathbf{0}, \qquad \forall j \in J, \tag{8c}$$

$$\lambda_j \left(\left[X_j^R + \sum_{l=1}^{L} \rho_{jl} P_l^R \right] - X^R \right) \leq \mathbf{0}, \qquad \forall j \in J, \tag{8d}$$

$$\sum_{j=1}^{n} \lambda_j = 1, \tag{8e}$$

$$X^R \leq \bar{X}^R, \tag{8f}$$

$$Y^R \leq \bar{Y}^R, \tag{8g}$$

$$\lambda, \pi \geq \mathbf{0}, \rho_{jl} \geq 0, \ \forall j, l. \tag{8h}$$

Let us consider the meaning of conditions (8). Their structure is similar to the structure of conditions (4) defining the R-VRS technology $T^{\mathrm{R}}_{\mathrm{VRS}}$ of Olesen et al. (2015).

First, assume that all inputs and outputs are volume measures. In this case, we remove inequalities (8c), (8d), (8f) and (8g) and all scalars ρ_{jl} from the statement (8). The resulting technology is the VRS technology with production trade-offs (5) of Podinovski (2004). In this case, the convex combinations of the observed DMUs are described by the first sums (taken with the weights λ_j) on the left-hand side of inequalities (8a) and (8b). These convex combinations are subsequently modified by the application of trade-offs (5) taken in proportions π_t, $t = 1,\dots,K$. As assumed by Axiom 4, such modifications keep the resulting DMU in the technology.

Let us now consider the general case of conditions (8). In this case, the ratio trade-offs (6) are used in conditions (8c) and (8d) to modify ratio inputs and outputs of individual observed DMUs. This is reflected by the fact that the proportions ρ_{jl} depend both on the trade-off l and the observed DMU j.

We can now explain the meaning of conditions (8). The inequalities (8a) and (8b) mean that the convex combinations of the vectors X_j^V and Y_j^V of volume inputs and outputs of the observed DMUs (X_j, Y_j), $j \in J$, which are further modified by production trade-offs (5) taken in proportions π_t, $t = 1,\dots,K$, outperform the vectors of volume measures X^V and Y^V of the DMU (X,Y).

If we do not have ratio trade-offs (6), the inequalities (8c) and (8d) become (4c) and (4d). The latter mean that the ratio inputs and outputs of the observed DMUs (X_j, Y_j) that enter the convex combinations in (8a) and (8b) with a $\lambda_j > 0$, are not worse than the corresponding ratio measures of the DMU (X,Y). In a typical practical application, this means that every observed DMU j with a positive λ_j operates in an environment, which is not more favourable (less harsh) than the environment of DMU (X,Y), and the quality of its outputs is not lower than the quality of outputs of DMU (X,Y).

If the ratio trade-offs (6) are specified, then a similar interpretation remains valid, with an additional step. Namely, as seen from inequalities (8c) and (8d), the ratio inputs and outputs of the observed DMUs are first adjusted by the ratio trade-offs (6). If the ratio measures of the observed DMU (X_j, Y_j) can be modified by the ratio trade-offs (6) in such a way that they are not worse than the ratio measures of the DMU (X,Y), then such modified observed DMU may enter the convex combination of the volume measures in (8a) and (8b) with a $\lambda_j > 0$. Otherwise, such observed DMU has a zero weight λ_j in the convex combination of the volume measures.

It is clear that the original R-VRS technology $T^{\mathrm{R}}_{\mathrm{VRS}}$ of Olesen et al. (2015) is a subset of technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$, i.e., we have the following embedding:

$$T^{\mathrm{R}}_{\mathrm{VRS}} \subseteq T^{\mathrm{R}}_{\mathrm{VRS-TO}}. \tag{9}$$

To see this, note that any DMU $(X,Y) \in T^{\mathrm{R}}_{\mathrm{VRS}}$ satisfies conditions (4) with some vector $\lambda \in \mathbb{R}^n$. Then the DMU (X,Y) satisfies conditions (8) with the same vector λ and all scalars π_t and ρ_{jl} taken equal to zero, for all t, j, l. By Theorem 1, DMU $(X,Y) \in T^{\mathrm{R}}_{\mathrm{VRS-TO}}$, and the embedding (9) follows.

This result means that the efficiency of any DMU assessed in technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ cannot be higher (and, as shown by an application in Section 7, is often lower) than its efficiency in technology $T^{\mathrm{R}}_{\mathrm{VRS}}$. In other words, the specification of production trade-offs generally leads to improved discrimination on efficiency.

We now obtain two useful properties of technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$, which are generalizations of similar properties of technology $T^{\mathrm{R}}_{\mathrm{VRS}}$ proved by Olesen, Petersen, & Podinovski (2022b).

Theorem 2. *Technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ is the union of a finite number of polyhedral sets.*

Corollary 1. *Technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ is a closed set.*

5. The F-CRS technology with production trade-offs

The conventional CRS technology may be viewed as the extension of the VRS technology in which we allow all DMUs to be scaled with a nonnegative scalar α. Olesen et al. (2015) develop a similar extension of the R-VRS technology, referred to as the Ratio-CRS (R-CRS) technology. In the R-CRS technology, the volume inputs and outputs are assumed scalable, as in the standard CRS technology. However, the ratio inputs and outputs may change or remain constant, according to one of the four types of ratio measures.

Of particular practical importance is the R-CRS technology in which the volume inputs and outputs are scalable, provided the ratio inputs and outputs do not change. This is the *fixed* type of ratio measures in the classification of Olesen et al. (2015). Such fixed ratios are typically used to capture environmental conditions (e.g., socio-economic characteristics of the production process) or quality of resources, goods and services provided. We use the fixed type of ratio measures to model the production process with fully scalable volume inputs and outputs, provided the environmental and quality characteristics of the production process are kept unchanged.

The scalability of volume inputs and outputs with the fixed ratio measures is stated by the following axiom.

Axiom 5 Scalability of volume inputs and outputs**.** Let $(X^V, X^R, Y^V, Y^R) \in T$. Then, for all $\alpha \geq 0$, $(\alpha X^V, X^R, \alpha Y^V, Y^R) \in T$.

Olesen et al. (2015) refer to the R-CRS technology in which all ratio inputs and outputs are of the fixed type as the F-CRS technology and denote it $T^{\mathrm{F}}_{\mathrm{CRS}}$. They obtain three explicit operational statements of the F-CRS technology. In order to avoid lengthy repetitions, we do not reproduce these statements and develop a standalone extension of the F-CRS technology by production trade-offs. The original F-CRS technology is a special case of this extension obtained by simply omitting all production trade-offs from the model.

We start by defining the extension of technology $T^{\mathrm{F}}_{\mathrm{CRS}}$ obtained by the incorporation of production trade-offs (5) and (6), which we denote $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$.

Definition 2. Technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ is the intersection of all sets $T \subset \mathbb{R}^m_+ \times \mathbb{R}^s_+$ that satisfy Axioms 1–5.

Similar to the case of R-VRS, it is straightforward to prove that technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ satisfies all Axioms 1–5 and can therefore be viewed as the smallest among all technologies that satisfy these axioms.

5.1. A nonlinear statement

The following theorem provides an equivalent explicit statement of technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$. It can be regarded as an extension (allowing additional production trade-offs) of Theorem 2 of Olesen et al. (2015) for the case in which all ratio measures are of the fixed type.

Theorem 3. *Technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ is the set of all DMUs $(X, Y) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ for which there exist vectors $\lambda, \alpha \in \mathbb{R}^n$, $\pi \in \mathbb{R}^K$ and scalars ρ_{jl}, $j \in J$, $l = 1, \ldots, L$, such that*

$$\sum_{j=1}^{n} \lambda_j \alpha_j Y_j^V + \sum_{t=1}^{K} \pi_t Q_t^V \geq Y^V, \tag{10a}$$

$$\sum_{j=1}^{n} \lambda_j \alpha_j X_j^V + \sum_{t=1}^{K} \pi_t P_t^V \leq X^V, \tag{10b}$$

$$\lambda_j \left(\left[Y_j^R + \sum_{l=1}^{L} \rho_{jl} Q_l^R \right] - Y^R \right) \geq \mathbf{0}, \qquad \forall j \in J, \tag{10c}$$

$$\lambda_j \left(\left[X_j^R + \sum_{l=1}^{L} \rho_{jl} P_l^R \right] - X^R \right) \leq \mathbf{0}, \qquad \forall j \in J, \tag{10d}$$

$$\sum_{j=1}^{n} \lambda_j = 1, \tag{10e}$$

$$X^R \leq \bar{X}^R, \tag{10f}$$

$$Y^R \leq \bar{Y}^R, \tag{10g}$$

$$\lambda, \alpha, \pi \geq \mathbf{0}, \ \rho_{jl} \geq 0, \ \forall j, l. \tag{10h}$$

The meaning of the above conditions is similar to the meaning of the corresponding conditions in the statement (8) of technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$. The obvious difference is the specification of the scaling factors α_j, $j \in J$, in the statement (10). In line with Axiom 5, these factors allow proportional scaling of the volume inputs and outputs of the observed DMUs, without changing their ratio measures.

Let us make two useful remarks. First, the original F-CRS technology $T^{\mathrm{F}}_{\mathrm{CRS}}$ of Olesen et al. (2015) is a subset of technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$, i.e.,

$$T^{\mathrm{F}}_{\mathrm{CRS}} \subseteq T^{\mathrm{F}}_{\mathrm{CRS-TO}}. \tag{11}$$

(The proof follows from the fact that the statement of technology $T^{\mathrm{F}}_{\mathrm{CRS}}$ corresponds to the special case of the statement (10) of technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ in which all scalars π_t and ρ_{jl} are taken equal to zero.)

The embedding (11) means that the efficiency of any DMU (X, Y) in technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ is not higher than in technology $T^{\mathrm{F}}_{\mathrm{CRS}}$. Therefore, similar to the case of R-VRS, the use of production trade-offs in the case of F-CRS generally results in improved discrimination on efficiency.

Second, it is clear that, if we consider technologies $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ and $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ generated by the same set of observed DMUs and incorporating the same set of production trade-offs (5) and (6), then the former technology is a subset of the latter, i.e., we always have:

$$T^{\mathrm{R}}_{\mathrm{VRS-TO}} \subseteq T^{\mathrm{F}}_{\mathrm{CRS-TO}}. \tag{12}$$

Indeed, according to Theorem 1, any DMU $(X, Y) \in T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ satisfies conditions (8), together with some vector $\lambda \in \mathbb{R}^n$ and scalars π_t, $t = 1, \ldots, K$, and ρ_{jl}, $j \in J$, $l = 1, \ldots, L$. Then this DMU also satisfies all conditions (10) with the same parameters λ, π_t and ρ_{jl}, if we take all scalars $\alpha_j = 1$, for all $j \in J$. By Theorem 3, $(X, Y) \in T^{\mathrm{F}}_{\mathrm{CRS-TO}}$, and the embedding (12) follows.

5.2. A partly linearized statement

We now partly linearize the statement (10) of technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$, by adapting the approach of Olesen et al. (2015). In this approach, we introduce three nonnegative vectors $\kappa, \mu, \nu \in \mathbb{R}^n_+$ and make the substitution $\lambda_j \alpha_j = \kappa_j - \nu_j + \mu_j$, for all $j \in J$, which explains the new conditions (13a) and (13b). As the proof of the next theorem shows, this substitution requires some further changes to the statement of the technology.

Theorem 4. *Technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ is the set of all DMUs $(X, Y) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ for which there exist vectors $\kappa, \mu, \nu \in \mathbb{R}^n$, $\pi \in \mathbb{R}^K$ and scalars ρ_{jl}, $j \in J$, $l = 1, \ldots, L$, such that*

$$\sum_{j=1}^{n} (\kappa_j + \mu_j - \nu_j) Y_j^V + \sum_{t=1}^{K} \pi_t Q_t^V \geq Y^V, \tag{13a}$$

$$\sum_{j=1}^{n} (\kappa_j + \mu_j - \nu_j) X_j^V + \sum_{t=1}^{K} \pi_t P_t^V \leq X^V, \tag{13b}$$

$$(\kappa_j + \mu_j) \left(\left[Y_j^R + \sum_{l=1}^{L} \rho_{jl} Q_l^R \right] - Y^R \right) \geq \mathbf{0}, \qquad \forall j \in J, \tag{13c}$$

$$(\kappa_j + \mu_j) \left(\left[X_j^R + \sum_{l=1}^{L} \rho_{jl} P_l^R \right] - X^R \right) \leq \mathbf{0}, \qquad \forall j \in J, \tag{13d}$$

$$\sum_{j=1}^{n} \kappa_j = 1, \tag{13e}$$

$$\kappa_j - \nu_j \geq 0, \qquad \forall j \in J, \tag{13f}$$

$$X^R \leq \bar{X}^R, \tag{13g}$$

$$Y^R \leq \bar{Y}^R, \tag{13h}$$

$$\kappa, \mu, \nu, \pi \geq \mathbf{0}, \rho_{jl} \geq 0, \ \forall j, l. \tag{13i}$$

Theorem 4 states technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ in an "almost" linear form. (The nonlinear conditions (13c) and (13d) are easy to linearize by using the "Big M" approach—see Section 6.) It also allows us to obtain the following properties of technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$, which represent a generalization of the results proved by Olesen et al. (2022b) for the F-CRS technology without trade-offs.

Theorem 5. *Technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ is the union of a finite number of polyhedral sets.*

Corollary 2. *Technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ is a closed set.*

5.3. An imperfect simplified statement

Theorems 3 and 4 present two alternative but equivalent statements (10) and (13) of technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$. The former is useful for exploring axiomatic properties of technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ and for explaining the meaning of DMUs in this technology. The latter is useful for establishing some further theoretical properties as in Theorem 5 and Corollary 2 and also for practical computations.

Below we obtain an alternative simplified statement of technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ in which only one intensity vector λ is used, instead of the three vectors κ, μ and ν as in statement (13). This statement has obvious computational advantages but should be used with caution. The potential problem is that the simplified statement defines a set that is slightly larger than the F-CRS technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$. Whether this affects the results of efficiency calculations can be established by a simple check of the optimal solution, as discussed below.

Consider the following technology which is defined by the same conditions (8) as the R-VRS technology from which the normalizing equality (8e) is removed. This definition may appear to be an intuitive extension of the R-VRS technology with production trade-offs to its F-CRS analogue. Surprisingly, as discussed below, this is not a perfect extension, although it *may* be good enough for practical computations.

Definition 3. Technology $\hat{T}^{\mathrm{F}}_{\mathrm{CRS-TO}}$ is the set of all DMUs $(X, Y) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ for which there exist vectors $\lambda \in \mathbb{R}^n$, $\pi \in \mathbb{R}^K$ and scalars ρ_{jl}, $j \in J$, $l = 1, \ldots, L$, such that

$$\sum_{j=1}^{n} \lambda_j Y_j^V + \sum_{t=1}^{K} \pi_t Q_t^V \geq Y^V, \tag{14a}$$

$$\sum_{j=1}^{n} \lambda_j X_j^V + \sum_{t=1}^{K} \pi_t P_t^V \leq X^V, \tag{14b}$$

$$\lambda_j \left(\left[Y_j^R + \sum_{l=1}^{L} \rho_{jl} Q_l^R \right] - Y^R \right) \geq \mathbf{0}, \qquad \forall j \in J, \tag{14c}$$

$$\lambda_j \left(\left[X_j^R + \sum_{l=1}^{L} \rho_{jl} P_l^R \right] - X^R \right) \leq \mathbf{0}, \qquad \forall j \in J, \tag{14d}$$

$$X^R \leq \bar{X}^R, \tag{14e}$$

$$Y^R \leq \bar{Y}^R, \tag{14f}$$

$$\lambda, \pi \geq \mathbf{0}, \rho_{jl} \geq 0, \ \forall j, l. \tag{14g}$$

Theorem 6. *We have $T^{\mathrm{F}}_{\mathrm{CRS-TO}} \subseteq \hat{T}^{\mathrm{F}}_{\mathrm{CRS-TO}}$. If DMU $(X, Y) \in \hat{T}^{\mathrm{F}}_{\mathrm{CRS-TO}}$ satisfies conditions (14) with some vectors λ, π and scalars ρ_{jl} such that $\lambda \neq \mathbf{0}$, then $(X, Y) \in T^{\mathrm{F}}_{\mathrm{CRS-TO}}$.*

According to Theorem 6, technology $\hat{T}^{\mathrm{F}}_{\mathrm{CRS-TO}}$ is slightly larger than technology $T^{\mathrm{F}}_{\mathrm{CRS-TO}}$. The reason of this discrepancy is that the statement (14) allows λ to be a zero vector, while this is disallowed by the statements (10) and (13). This in turn means that, if $\lambda = \mathbf{0}$, all inequalities for the ratio measures (14c) and (14d) are trivially satisfied for all $j \in J$, regardless of the vectors X^R and Y^R of the DMU (X, Y), while this is not so for the corresponding inequalities in the statements (10) and (13).

It is clear that the situations in which a DMU (X, Y) satisfies conditions (14) with a zero vector λ should be rare in practical applications. Indeed, if $\lambda = \mathbf{0}$, the inequalities (14a) and (14b) imply that the vectors Y^V and X^V are outperformed purely by a combination of trade-off vectors Q_t^V and P_t^V, taken with the weights π_t, $t = 1, \ldots, K$. This means that either the DMU (X, Y) satisfying conditions (14) with $\lambda = \mathbf{0}$ is extremely inefficient or that the production trade-offs are too demanding.

In the application discussed in Section 7, we originally used both the full and simplified statements (13) and (14). The results of thousands of computations using both statements were identical, except a few rare cases. In all such cases, the reason of discrepancy was that the optimal vector λ in the simplified model was a zero vector, and the results obtained by solving the simplified model were incorrect.

The lesson learned was that, if we use the simplified model of technology (14), we need to perform an additional check of the vectors λ in all optimal solutions. The case in which λ is a zero vector should be a rare occurrence, which may point to a problem with the data set or value judgements stated as production trade-offs. If both the data and assumed trade-offs are unproblematic but all components of the optimal vector λ are equal to zero, the only theoretically sound alternative is to perform calculations using statement (13). In our application in Section 7, we eventually performed all computations using statement (13).

Remark 2. Olesen et al. (2015) consider the full and simplified statements of the F-CRS technology without production trade-offs. In this case, the trade-off terms in the inequalities (14a) and (14b) do not appear, and the case $\lambda = \mathbf{0}$ implies that $Y^V = \mathbf{0}$. This case is clearly irrelevant for practical applications and the possibility of λ being a zero vector can be ignored. Therefore, if no production trade-offs are specified, the simplified model of technology can be used for all practical computations.

6. Solving R-VRS and F-CRS models with production trade-offs

In this section, we consider solving DEA models based on the R-VRS and F-CRS technologies expanded by production trade-offs. To be specific, we consider the assessment of output radial efficiency of the DMUs in the R-VRS technology and only briefly comment on the case of F-CRS technology afterwards. The case of input radial efficiency and other, including nonradial, measures is similar and is not discussed.

To unify the discussion and avoid the consideration of special cases, consider the assessment of output radial efficiency of DMU (X_o, Y_o) in technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ with respect to all, volume and ratio, outputs. In this case, we attach the output improvement factor η

to all outputs. (In practice, we may consider improvements to the volume and ratio outputs separately, as illustrated by the application in Section 7.)

The output radial efficiency of DMU (X_o, Y_o) is the inverse to the optimal value η^* of the output-oriented R-VRS program:

$$\eta^* = \max \quad \eta \tag{15a}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} \lambda_j Y_j^V + \sum_{t=1}^{K} \pi_t Q_t^V \geq \eta Y_o^V, \tag{15b}$$

$$\sum_{j=1}^{n} \lambda_j X_j^V + \sum_{t=1}^{K} \pi_t P_t^V \leq X_o^V, \tag{15c}$$

$$\lambda_j \left(\left[Y_j^R + \sum_{l=1}^{L} \rho_{jl} Q_l^R \right] - \eta Y_o^R \right) \geq \mathbf{0}, \qquad \forall j \in J, \tag{15d}$$

$$\lambda_j \left(\left[X_j^R + \sum_{l=1}^{L} \rho_{jl} P_l^R \right] - X_o^R \right) \leq \mathbf{0}, \qquad \forall j \in J, \tag{15e}$$

$$\sum_{j=1}^{n} \lambda_j = 1, \tag{15f}$$

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

$$\lambda, \pi \geq \mathbf{0}, \rho_{jl} \geq 0, \ \forall j, l; \ \ \eta \text{ sign-free.} \tag{15h}$$

Note that constraint (15g) specifies the upper bound on the improved subvector ηY_o^R. Also note that program (15) does not include inequality (8f) because the ratio input vector X_o^R of DMU (X_o, Y_o) does not change.

The nonlinear inequalities (15d) and (15e) of program (15) can be linearized using the well-known "Big M" approach. Namely, we first restate conditions (15d) and (15e) in the following "either-or" form:

$$\begin{aligned} &\text{either} \quad \lambda_j = 0 \\ &\text{or} \quad \left\{ \eta Y_o^R - \left[Y_j^R + \sum_{l=1}^{L} \rho_{jl} Q_l^R \right] \leq \mathbf{0} \right. \\ &\text{and} \left. \left[X_j^R + \sum_{l=1}^{L} \rho_{jl} P_l^R \right] - X_o^R \leq \mathbf{0} \right\}, \qquad \forall j \in J. \end{aligned} \tag{16}$$

Introducing binary variables δ_j, $j \in J$, we linearize conditional statements (16) as

$$\lambda_j \leq \delta_j, \qquad \forall j \in J, \tag{17a}$$

$$\eta Y_o^R - \left[Y_j^R + \sum_{l=1}^{L} \rho_{jl} Q_l^R \right] \leq L_1 (1 - \delta_j), \qquad \forall j \in J, \tag{17b}$$

$$\left[X_j^R + \sum_{l=1}^{L} \rho_{jl} P_l^R \right] - X_o^R \leq L_2 (1 - \delta_j), \qquad \forall j \in J, \tag{17c}$$

where the vectors $L_1 \in \mathbb{R}_+^{s^R}$ and $L_2 \in \mathbb{R}_+^{m^R}$ have sufficiently large positive components.

Following the "Big M" approach, the components of vectors L_1 and L_2 should be so large that the inequalities (17b) and (17c) become redundant in the case $\delta_j = 0$. Details of this approach can be found in Olesen, Petersen, & Podinovski (2017).

Remark 3. The described linearization approach is also applicable in the case of F-CRS technology with trade-offs, whose full statement is given by conditions (13) or, in a simplified form, by conditions (14). In the latter case, we solve program (15) from which the normalizing equality (15f) is removed. Following the "Big M" approach, we introduce binary variables δ_j, $j \in J$, and replace the nonlinear inequalities (15d) and (15e) by the inequalities (17), in which the inequality (17a) requires the following correction. Namely, because the variables λ_j are no longer bounded above by 1 as in the case of R-VRS, the inequality (17a) is replaced by

$$\lambda_j \leq M_1 \delta_j, \qquad \forall j \in J,$$

with a sufficiently large constant $M_1 > 0$. If we solve the output-oriented program based on the full statement of technology (13), we replace the inequality (17a) by the conditions

$$\lambda_j + \mu_j \leq M_2 \delta_j, \qquad \forall j \in J,$$

with a sufficiently large constant $M_2 > 0$.

Remark 4. The authors are grateful to the anonymous reviewer who pointed out that, as an alternative to the "Big M" method, the nonlinear conditions (15d) and (15e) may be restated using the special ordered sets (SOS) of Beale & Tomlin (1970). Exploring this possibility is left outside the scope of our paper for future research.

7. Application to secondary schools in England

In this section, we illustrate the methodology developed in our paper by its application to the assessment of efficiency of a large sample of secondary schools in England. In this country, pupils enter secondary education at the age of 11. They are expected to take national exams for General Certificate of Secondary Education (GCSE) at the age of 16. After completing GCSEs, many pupils proceed to the sixth form to obtain A-level qualifications, typically at the age of 18, used as admissions criteria by the universities.

7.1. Data

For homogeneity reasons, we consider only the secondary schools classed as academies and free schools, collectively referred to as *academies*. Approximately 80% of secondary schools in England are academies, covering 79% of secondary school pupils (DfE, 2022). We further limit the sample only to academies with the sixth forms and non-selective admissions policies. We also exclude London schools because of the known "London effect", i.e., taking into account that London schools typically outperform schools in the rest of England, especially among disadvantaged areas and pupils (Ross, Lessof, Brind, Khandker, & Aitken, 2020).

The final sample for this study includes 891 academies. All data were collected in the academic year 2018–2019 and provided to us by the Department for Education. Table 1 shows descriptive statistics for the four inputs and three outputs used in this application.

Following the literature on school efficiency (see, e.g., Brennan, Haelermans, & Ruggiero, 2014, and Silva, Camanho, & Barbosa, 2020), we use *volume inputs* 1 and 2 to account for the number of teachers and school expenditure (excluding teacher salaries), respectively. We also include two *ratio inputs*. Namely, input 3 represents the percentage of pupils with middle or high prior attainment at the beginning of secondary education. Input 4 shows the percentage of pupils not receiving free school meals. Both ratio inputs are commonly considered as having a positive impact on the attainment of school leavers and are often used in reported applications (Bradley, Johnes, & Millington, 2001; Brennan et al., 2014; Thanassoulis & Dunstan, 1994).

We consider the total number of pupils in all years as the single *volume output* 1 (see Remark 5). The *ratio output* 2 represents

Table 1
Descriptive statistics for 891 academies in application.

Inputs and outputs	Mean	Standard deviation	Minimum	Maximum
Input 1: Teachers	73.77	22.43	20.5	155.1
Input 2: Expenditure (thousand pounds)	3,053	1,047.03	326	9,244
Input 3: Good attainment on entry (%)	89.57	6.34	0	100
Input 4: No free school meals (%)	88.18	7.99	47.9	99.2
Output 1: Pupils	1,205.65	359.25	204	2,500
Output 2: Good GCSEs (%)	43.2	14.14	0	89
Output 3: Top universities (%)	20	13.34	0	75

the percentage of pupils achieving strong passes in both GCSE English and Mathematics. The *ratio output* 3 shows the percentage of pupils admitted to the top third of universities. These ratio outputs are included as quality characteristics of the two different stages of the education process and are calculated according to the established methodology of the Department for Education.

Remark 5. The treatment of the number of pupils as an output and not as an input is consistent with the assumed Axiom 2 of free disposability. Namely, for the given input levels, any school should, if required, be able to teach fewer students but not more. The latter would be assumed by the model if pupils were treated as an input, which is clearly problematic.

Another consideration is the association of larger pupil numbers with higher efficiency of the schools. To illustrate this, assume that we have two schools A and B that have identical number of teachers and expenditures, and whose ratio inputs and outputs are also the same. Suppose that the only difference is that school A teaches 500 pupils and school B teaches 1000 pupils. We regard school B as more efficient than school A, which is consistent with the treatment of the number of pupils as an output.

7.2. Identification of production trade-offs

We use seven production trade-offs that reflect the assumed relationship between different inputs and outputs. Below, we state these trade-offs formally and explain their meaning.

Judgement 1. If required, any school should be able to increase its number of pupils by 10, provided the school employs 1 extra teacher and its budget is increased by £20,000, while the remaining inputs and outputs are kept unchanged.

In line with notation (5), and taking into account that data on expenditure is represented in thousands of pounds, the above judgement is stated as the following production trade-off:

$$P_1^V = (1, 20)^\top, Q_1^V = (10).$$

The stated trade-off is deemed to be a sufficiently conservative judgement which all schools (especially the efficient ones that form the frontier against which all other schools are benchmarked) should find unproblematic. For example, in our data set, the pupil-to-teacher ratio for individual schools ranges between 7 and 33, and the assumed trade-off, which requires an increase of the number of teachers and the funding, appears to be a sufficient compensation to the school for the increase of its pupil cohort by 10.

It is also worth highlighting that the simultaneous change stated by this trade-off assumes that the ratio inputs and outputs remain unchanged. In other words, the assumed change of the volume measures is possible without any change to the socio-economic and quality characteristics of the school cohorts and the teaching process.

Judgement 2. If required, the simultaneous reduction of the number of teachers by 1 and the number of pupils by 15 is technologically possible for any school, provided the remaining inputs and outputs are kept unchanged.

We can restate the above judgement as follows:

$$P_2^V = (-1, 0)^\top, Q_2^V = (-15).$$

As a variant, we could modify the above judgement by including a simultaneous reduction of the budget of the school. However, such trade-off would be more demanding because, after its application, the resulting school would be assumed technologically possible with both the reduced number of teachers and, additionally, the reduced budget. To keep our judgements more conservative, we do not require any reduction of the budget. (Note that, in Judgement 1, we include an increase of the budget for exactly the same reason of making the trade-off reasonably conservative.)

Judgement 3. A loss of 1 teacher at any school can be compensated by the additional budget of £50,000 (which, for example, could be used to pay for an external substitution teacher), assuming the remaining inputs and outputs are kept unchanged.

Judgement 3 is formally stated as:

$$P_3^V = (-1, 50)^\top, Q_3^V = (0).$$

We now proceed to stating trade-offs between ratio measures. The underlying judgements take into account the fact that prior achievement and socio-economic background of pupils can almost completely explain academic differences between schools—see, e.g., Gorard (2014).

Judgement 4. If the percentage of pupils with good attainment levels on entry to school increases by 1%, then the percentage of pupils achieving good GCSE results should go up by at least 0.5% and the percentage of pupils proceeding to the top third of universities should increase by at least 0.25%, provided the remaining inputs and outputs are kept unchanged.

We present the above judgement as the following trade-off:

$$P_1^R = (1, 0)^\top, Q_1^R = (0.5, 0.25)^\top.$$

Judgement 5. If the percentage of pupils with good attainment levels on entry to school is reduced by 1%, then the percentage of pupils achieving good GCSE results and those going to top universities should not decline by more than the same percentage, i.e., by 1% each, provided the remaining inputs and outputs are kept unchanged.

We state this judgement as:

$$P_2^R = (-1, 0)^\top, Q_2^R = (-1, -1)^\top.$$

The final two judgements represent the assumption that the percentage of pupils not receiving free school meals has a similar impact on the quality of the teaching process as the percentage of pupils with good attainment on entry to school.

Judgement 6. If the percentage of pupils not receiving free school meals increases by 1%, then the percentage of pupils achieving good GCSE results should go up by at least 0.5% and the percentage of pupils proceeding to the top third of universities should increase by at least 0.25%, provided the remaining inputs and outputs are kept unchanged.

Table 2
Efficiency with respect to the volume output in the R-VRS models.

Statistics	No trade-offs	J1	J1–J2	J1–J3	J1–J4	J1–J5	J1–J6	J1–J7
Average efficiency (%)	94.23	93.76	93.19	91.53	90.81	89.04	88.50	82.36
Minimum efficiency (%)	58.44	58.44	41.25	25.02	24.89	24.89	24.89	24.89
Number of efficient schools	443	412	393	343	309	231	209	96
% of inefficient schools	50.28	53.76	55.89	61.50	65.32	74.07	76.54	89.23

Table 3
Efficiency with respect to the volume output in the F-CRS models.

Statistics	No trade-offs	J1	J1–J2	J1–J3	J1–J4	J1–J5	J1–J6	J1–J7
Average efficiency (%)	90.45	90.35	89.34	87.47	86.54	84.47	83.65	74.72
Minimum efficiency (%)	40.97	40.97	18.79	17.5	17.5	17.5	17.5	17.5
Number of efficient schools	299	290	285	245	225	146	131	35
% of inefficient schools	66.44	67.45	68.01	72.50	74.86	83.61	85.30	96.07

Table 4
Efficiency with respect to the vector of ratio outputs in the R-VRS models.

Statistics	No trade-offs	J1	J1–J2	J1–J3	J1–J4	J1–J5	J1–J6	J1–J7
Average efficiency (%)	89.4	88.25	87.3	85.6	83.93	80.54	78.08	70.35
Minimum efficiency (%)	15.91	15.91	15.91	15.91	15.56	15.56	12.12	12.12
Number of efficient schools	472	446	420	368	318	243	209	97
% of inefficient school	47.03	49.94	52.86	58.70	64.31	72.73	76.54	89.11

We convert the above judgement to the following trade-off:

$P_3^R = (0, 1)^\top, Q_3^R = (0.5, 0.25)^\top.$

Judgement 7. If the percentage of pupils not receiving free school meals is reduced by 1%, then the percentage of pupils achieving good GCSE results and those going to top universities should not decline by more than the same percentage, i.e., by 1% each, provided the remaining inputs and outputs are kept unchanged.

This judgement is stated as follows:

$P_4^R = (0, -1)^\top, Q_4^R = (-1, -1)^\top.$

7.3. Efficiency with respect to the volume output

We first consider the following question: what is the maximum number of pupils that a school can teach, for the given level of its resources (teachers and expenditure) and assuming that the socio-economic and quality characteristics of the teaching process (represented by the two ratio inputs and two ratio outputs) do not change? This question should be of interest to local education authorities who may consider allocating additional pupils to the schools.

We explore this question in the R-VRS and F-CRS technologies. In both cases, we first consider the technology without trade-offs and then the technologies obtained from it by consecutive incorporation of the trade-offs representing Judgements 1–7, as discussed in Section 7.2. The general statement of all such R-VRS models is given by program (15) in which we keep the improvement factor η attached to the volume output vector Y_o^V in constraints (15a) and remove η from constraints (15d). In the case of F-CRS, we solve similar programs based on its statement (13).

Tables 2 and 3 present a summary of computational results. (We convert the efficiency scores in the range [0,1] obtained by solving the R-VRS and F-CRS models to percentages and round the results to two decimal places.) Their second columns correspond to the models solved without trade-offs. The remaining columns correspond to the models in which we progressively incorporate additional trade-offs based on Judgements 1–7, denoted J1–J7. For example, the columns "J1" correspond to the models in which we use the single trade-off based on Judgement 1. The columns "J1–J2" correspond to the use of the trade-offs based on Judgements 1 and 2. The last column corresponds to the use of all seven production trade-offs.

In line with the theoretical embeddings (9) and (11), the two tables show that the incorporation of production trade-offs has a significant impact on the discriminating power of the model. Thus, the average efficiency across all schools, as assessed by the standard R-VRS model of Olesen et al. (2017) used without trade-offs, is 94.33%, and just over one half (50.28%) of all schools are inefficient. The incorporation of production trade-offs gradually improves discrimination. The final R-VRS model with all seven trade-offs reduces the average efficiency to 82.36% and identifies inefficiency in 89.23% of all schools. A similar pattern is seen in the case of F-CRS model, in which the average efficiency is reduced from 90.45% to 74.72%, and the number of inefficient schools increases from 66.44% to 96.07%.

We also note that, in line with the embedding (12), the efficiency of schools evaluated in the F-CRS technology used with any set of trade-offs is generally lower than in the R-VRS technology used with the same trade-offs.

7.4. Efficiency with respect to the ratio outputs

We now consider a different question: what is the maximum proportional increase to the two ratio outputs representing the quality of the teaching process (percentage of pupils achieving good GCSE results and proceeding to top universities) that the school should be able to achieve, given its resources (teachers and expenditure), the number of pupils and the socio-economic characteristics of the school cohorts, represented by the two ratio inputs. This assessment scenario is particularly relevant when, out of all volume and ratio measures, only the two ratio outputs representing quality of education are discretionary, while the other are exogenous and are not under the school's control.

To answer the stated question in the R-VRS technology with different sets of trade-offs, we solve program (15) in which we keep the improvement factor in constraints (15d) but remove it from constraints (15a). In the case of F-CRS technology, we use similar programs based on its statement (13).

Tables 4 and 5 present a summary of computational results in this scenario. Similar to the previous case, it is clear that the incorporation of production trade-offs in the R-VRS and F-CRS technolo-

Table 5
Efficiency with respect to the vector of ratio outputs in the F-CRS models.

Statistics	No trade-offs	J1	J1–J2	J1–J3	J1–J4	J1–J5	J1–J6	J1–J7
Average efficiency (%)	83.57	83.25	82.93	81.62	80.09	76.44	73.88	65.81
Minimum efficiency (%)	15.91	15.91	15.91	15.91	15.56	15.56	12.12	12.12
Number of efficient schools	316	309	300	278	244	171	137	39
% of inefficient schools	64.53	65.32	66.33	68.8	72.62	80.81	84.62	95.62

gies leads to significantly improved efficiency discrimination of the resulting models.

8. Conclusion

The use of value judgements in the form of weight restrictions and dual to them production trade-offs has been reported in many applications of the conventional VRS and CRS models. Such judgements allow the analyst to obtain a meaningful extension of the underlying technology, by specifying that certain simultaneous changes to the inputs and outputs should be possible for any DMU. This creates additional potential comparators for the DMUs and generally results in improved discrimination on efficiency.

In this paper, we introduce a similar notion of production trade-offs to the R-VRS and R-CRS models developed by Olesen et al. (2015) that allow the inputs and outputs to be specified by either volume or ratio measures. In the case of R-CRS, we focus on the most common situation in which the ratio inputs and outputs represent contextual (e.g., socio-economic) and quality characteristics of the production process. Such ratio measures remain constant while allowing the volume inputs and outputs to be scaled as in the standard CRS technology. We refer to such R-CRS technology with fixed ratio measures as the F-CRS technology.

In contrast with the conventional VRS and CRS models which can be stated as both primal and dual (envelopment and multiplier) linear programs, the R-VRS and F-CRS models are nonlinear and have only the envelopment form. Therefore, the value judgements traditionally stated in alternative but equivalent forms of weight restrictions and production trade-offs can now be stated only using the latter (trade-off) interpretation.

We define extensions of the R-VRS and F-CRS technologies by production trade-offs using the axiomatic approach. This means that every hypothetical DMU in the new technology can be explained by the explicitly stated assumptions about the technology. Namely, every DMU is either included in the R-VRS or F-CRS technology of Olesen et al. (2015) or is obtained from one of DMUs in the respective original technology by the application of the assumed production trade-offs. This in turn means that the target DMUs (e.g., radial targets) of inefficient DMUs are producible and, in line with the assumed axioms, should be regarded as valid benchmarks.

We distinguish between the production trade-offs specified for the volume inputs and outputs and production trade-offs involving changes to ratio measures. The mathematical approach to the incorporation of the former is similar to their use in the standard VRS and CRS models. However, the incorporation of production trade-offs between ratio measures requires a different mathematical approach.

We use an application in the context of school education to discuss the practical meaning of production trade-offs. As expected from theory, computational results confirm that the incorporation of production trade-offs in the R-VRS and F-CRS models results in improved discrimination on efficiency between the schools. It is clear that the proposed models with production trade-offs should provide similar advantages in applications in other contexts and sectors.

Our paper opens up further research avenues that have already been extensively explored in the case of conventional VRS and CRS models, and, more recently, in the R-VRS and R-CRS technologies. This includes exploring the notion of returns to scale and scale efficiency in the R-VRS technology (Olesen, Petersen, & Podinovski, 2022a) and issues of inconsistent trade-offs resulting in the introduction of "free production" in the VRS and CRS technologies (Podinovski & Bouzdine-Chameeva, 2013). Investigation of these issues in the R-VRS and F-CRS technologies expanded by production trade-offs is left for future research.

Acknowledgments

The authors are grateful to the Department for Education for providing the data set used in the application.

Appendix A. Proofs

Proof of Theorem 1. Denote T^* the set of all DMUs $(X, Y) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ for which there exists a vector $\lambda \in \mathbb{R}^n$ and scalars π_t and ρ_{jl}, $\forall j, l, t$, such that conditions (8) are true. We need to prove that $T^{\mathrm{R}}_{\mathrm{VRS-TO}} = T^*$. We first prove that T^* satisfies Axioms 1–4. Axioms 1 and 2 are straightforward. For example, Axiom 2 is true because, if DMU (X, Y) satisfies (8) with some scalars λ_j, π_t, ρ_{jl}, then $(\tilde{X}, \tilde{Y})$ also satisfies (8) with the same scalars. Lemmas 1 and 2 establish that T^* satisfies Axioms 3 and 4. Therefore, $T^{\mathrm{R}}_{\mathrm{VRS-TO}} \subseteq T^*$. The opposite embedding $T^* \subseteq T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ is established by Lemma 4. Therefore, $T^{\mathrm{R}}_{\mathrm{VRS-TO}} = T^*$. □

Lemma 1. *Technology T^* satisfies Axiom 3.*

Proof of Lemma 1. To prove that T^* satisfies Axiom 3, consider any two DMUs $(\tilde{X}, \tilde{Y}) \in T^*$ and $(\hat{X}, \hat{Y}) \in T^*$. These DMUs satisfy conditions (8) with some vectors $\tilde{\lambda}$ and $\hat{\lambda}$ and scalars $\tilde{\pi}_t$, $\tilde{\rho}_{jl}$ and $\hat{\pi}_t$ and $\hat{\rho}_{jl}$, $\forall j, t, l$, respectively. Consider any $\gamma \in [0, 1]$. Define

$$\left(\bar{X}^V, \bar{X}^R, \bar{Y}^V, \bar{Y}^R\right) = \gamma\left(\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R\right) + (1-\gamma)\left(\hat{X}^V, \hat{X}^R, \hat{Y}^V, \hat{Y}^R\right). \tag{A.1}$$

We need to prove that $(\bar{X}^V, \bar{X}^R, \bar{Y}^V, \bar{Y}^R) \in T^*$, i.e., that it satisfies (8) with some vector $\bar{\lambda}$ and scalars $\bar{\pi}_t$, $\bar{\rho}_{jl}$, $\forall j, t, l$. Define

$$\bar{\lambda} = \gamma\tilde{\lambda} + (1-\gamma)\hat{\lambda}, \qquad \bar{\pi} = \gamma\tilde{\pi} + (1-\gamma)\hat{\pi}. \tag{A.2}$$

To define multipliers $\bar{\rho}_{jl}$, first define the sets

$$\tilde{J} = \left\{j \in J \mid \tilde{\lambda}_j > 0\right\}, \qquad \hat{J} = \left\{j \in J \mid \hat{\lambda}_j > 0\right\}. \tag{A.3}$$

For each $l = 1, \ldots, L$, let:

$$\bar{\rho}_{jl} = \begin{cases} \tilde{\rho}_{jl}, & j \in \tilde{J}, \\ \hat{\rho}_{jl}, & j \in \hat{J} \setminus \tilde{J}, \\ 0, & \text{otherwise.} \end{cases} \tag{A.4}$$

Let us prove that the DMU $(\bar{X}^V, \bar{X}^R, \bar{Y}^V, \bar{Y}^R)$ defined by (A.1) satisfies all conditions (8) with the vectors $\bar{\lambda}$ and $\bar{\pi}$ defined by (A.2) and scalars $\bar{\rho}_{jl}$ defined by (A.4).

To prove (8a) and (8b), we state these conditions twice, for the DMU $(\tilde{X}, \tilde{Y})$ and vectors $\tilde{\lambda}$ and $\tilde{\pi}$, and for the DMU $(\hat{X}, \hat{Y})$ and vectors $\hat{\lambda}$ and $\hat{\pi}$. The proof is finalized by multiplying these inequalities by γ and $1-\gamma$, respectively, adding and rearranging the terms.

Let us prove (8c). It suffices to assume that $\bar{\lambda}_j > 0$. By (A.2), either $\tilde{\lambda}_j > 0$ or $\hat{\lambda}_j > 0$, or both. Then either $j \in \tilde{J}$ or $j \in \hat{J} \setminus \tilde{J}$. To be specific, let $j \in \tilde{J}$. Because DMU $(\tilde{X}, \tilde{Y})$ satisfies (8) with $\tilde{\lambda}$ and scalars $\tilde{\rho}_{jl}$, and because $\tilde{\lambda}_j > 0$, (8c) implies

$$\left[Y_j^R + \sum_{l=1}^{L} \tilde{\rho}_{jl} Q_l^R \right] - \tilde{Y}^R \geq \mathbf{0}. \tag{A.5}$$

Taking into account (3) and (A.1), we replace $\tilde{Y}^R$ by $\bar{Y}^R$. Also, by (A.4), because $j \in \tilde{J}$, we replace $\tilde{\rho}_{jl}$ by $\bar{\rho}_{jl}$, $\forall l$. With these replacements, and because $\bar{\lambda}_j > 0$, (A.5) implies (8c) for any $j \in \tilde{J}$. The proof for $j \in \hat{J} \setminus \tilde{J}$ is similar. The proof of conditions (8d) is also similar. The proof of conditions (8e)–(8g) is straightforward and is omitted. □

Lemma 2. *Technology T^*satisfies Axiom 4.*

Proof of Lemma 2. Consider any DMU $(X, Y) \in T^*$, and the DMU $(\tilde{X}, \tilde{Y})$ obtained from it as specified by Axiom 4. We need to prove that $(\tilde{X}, \tilde{Y}) \in T^*$. DMU (X, Y) satisfies conditions (8) with some vector $\hat{\lambda}$ and scalars $\hat{\pi}_l$, $\hat{\rho}_{jt}$. Let us prove that the DMU $(\tilde{X}, \tilde{Y})$ stated by (7) satisfies conditions (8) with the vector $\tilde{\lambda} = \hat{\lambda}$ and the scalars $\tilde{\pi}_t = \pi_t + \hat{\pi}_t$ and $\tilde{\rho}_{jl} = \rho_{jl} + \hat{\rho}_{jl}$, $\forall j, l, t$. The proof that conditions (8e)–(8h) are satisfied follows from the definition of $\tilde{\lambda}$, $\tilde{\pi}_t$ and $\tilde{\rho}_{jl}$ and the assumptions of Axiom 4. Let us prove conditions (8a)–(8d).

Consider condition (8a) stated for DMU (X, Y):

$$\sum_{j=1}^{n} \hat{\lambda}_j Y_j^V + \sum_{t=1}^{K} \hat{\pi}_t Q_t^V \geq Y^V. \tag{A.6}$$

Adding the sum $\sum_{t=1}^{K} \pi_t Q_t^V$ to both sides of (A.6), rearranging and noting the definition of $\tilde{Y}^V$ by (7), we have

$$\sum_{j=1}^{n} \tilde{\lambda}_j Y_j^V + \sum_{t=1}^{K} \tilde{\pi}_t Q_t^V \geq \tilde{Y}^V.$$

Consider conditions (8c) stated for DMU (X, Y). For any $j \in J$, we have

$$\hat{\lambda}_j \left(\left[Y_j^R + \sum_{l=1}^{L} \hat{\rho}_{jl} Q_l^R \right] - Y^R \right) \geq \mathbf{0}.$$

Adding and subtracting the sum $\sum_{l=1}^{L} \rho_{jl} Q_l^R$ inside the parentheses, rearranging the terms and using the definition of vector $\tilde{\lambda}$ and scalars $\tilde{\pi}_t$, we have

$$\tilde{\lambda}_j \left(\left[Y_j^R + \sum_{t=1}^{K} \tilde{\rho}_{jt} Q_t^R \right] - \tilde{Y}^R \right) \geq \mathbf{0}.$$

The proofs of conditions (8b) and (8d) is similar. Also, by the assumption of Axiom 4, $(\tilde{X}, \tilde{Y}) \in \mathbb{R}_+^m \times \mathbb{R}_+^s$. Therefore, $(\tilde{X}, \tilde{Y}) \in T^*$. □

Lemma 3. *Let technology T satisfy Axioms 2 and 4, and let $(X^V, X^R, Y^V, Y^R) \in T$. Define DMU $(\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R)$ by the following equalities, where scalars $\pi_t, \rho_l \geq 0$, $\forall t, l$, and S_Y^V, S_X^V, S_Y^R, $S_X^R \geq \mathbf{0}$ are slack vectors of appropriate dimensions:*

$$\tilde{Y}^V = Y^V + \sum_{t=1}^{K} \pi_t Q_t^V - S_Y^V, \tag{A.7a}$$

$$\tilde{X}^V = X^V + \sum_{t=1}^{K} \pi_t P_t^V + S_X^V, \tag{A.7b}$$

$$\tilde{Y}^R = Y^R + \sum_{l=1}^{L} \rho_l Q_l^R - S_Y^R, \tag{A.7c}$$

$$\tilde{X}^R = X^R + \sum_{l=1}^{L} \rho_l P_l^R + S_X^R. \tag{A.7d}$$

Further assume that $\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R \geq \mathbf{0}$ and that the subvectors $\tilde{X}^R$ and $\tilde{Y}^R$ are within the bounds (2). Then $\left(\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R\right) \in T$.

Proof of Lemma 3. It suffices to prove that the DMU $(\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R)$ can be obtained as a result of several consecutive modifications, starting with the DMU (X^V, X^R, Y^V, Y^R), by the trade-offs and slack variables such that all inputs and outputs of all intermediate DMU are nonnegative and within the bounds stated as (2). Then the proof is completed by noting Axioms 2 and 4.

Note that the DMU $\left(\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R\right)$ does not have to be computed in the stated order, i.e., applying the trade-offs in full proportions π_t and ρ_l first and subsequently adding or subtracting the full slack vectors. Instead, we can consecutively add the trade-offs and slack vectors in smaller proportions and in a different order, while keeping all intermediate DMUs nonnegative and satisfying the bounds (2). Different procedures implementing this idea can be suggested, but their precise description is tedious and is not given. □

Lemma 4. *The following embedding is true: $T^* \subseteq T^{\mathrm{R}}_{\mathrm{VRS-TO}}$.*

Proof of Lemma 4. Consider any DMU $(\tilde{X}, \tilde{Y}) \in T^*$. We need to prove that $(\tilde{X}, \tilde{Y}) \in T^{\mathrm{R}}_{\mathrm{VRS-TO}}$. DMU $(\tilde{X}, \tilde{Y})$ satisfies (8) with some vector $\tilde{\lambda}$ and scalars $\tilde{\pi}_t$ and $\tilde{\rho}_{jl}$, $\forall j, t, l$. Without loss of generality, let $\tilde{\lambda}_j > 0$, $\forall j \in J$. Then conditions (8) can be restated as equalities, in which the slack vectors $\tilde{S}_Y^V$, $\tilde{S}_X^V$, $\tilde{S}_{Yj}^R$, $\tilde{S}_{Xj}^R \geq \mathbf{0}$ are of appropriate dimensions:

$$\sum_{j=1}^{n} \tilde{\lambda}_j Y_j^V + \sum_{t=1}^{K} \tilde{\pi}_t Q_t^V - \tilde{S}_Y^V = \tilde{Y}^V, \tag{A.8a}$$

$$\sum_{j=1}^{n} \tilde{\lambda}_j X_j^V + \sum_{t=1}^{K} \tilde{\pi}_t P_t^V + \tilde{S}_X^V = \tilde{X}^V, \tag{A.8b}$$

$$Y_j^R + \sum_{l=1}^{L} \tilde{\rho}_{jl} Q_t^R - \tilde{S}_{Yj}^R = \tilde{Y}^R, \qquad \forall j \in J, \tag{A.8c}$$

$$X_j^R + \sum_{l=1}^{L} \tilde{\rho}_{jl} P_l^R + \tilde{S}_{Xj}^R = \tilde{X}^R, \qquad \forall j \in J, \tag{A.8d}$$

$$\sum_{j=1}^{n} \tilde{\lambda}_j = 1, \tag{A.8e}$$

$$\tilde{X}^R \leq \bar{X}^R, \tag{A.8f}$$

$$\tilde{Y}^R \leq \bar{Y}^R. \tag{A.8g}$$

For each observed DMU (X_j, Y_j), $j \in J$, conditions (A.8c) and (A.8d) are a special case of conditions (A.7c) and (A.7d). By Lemma 3, for each $j \in J$, we have $(X_j^V, \tilde{X}^R, Y_j^V, \tilde{Y}^R) \in T^{\mathrm{R}}_{\mathrm{VRS-TO}}$. By Axiom 3,

$$(X^V, X^R, Y^V, Y^R) = \left(\sum_{j=1}^{n} \tilde{\lambda}_j X_j^V, \tilde{X}^R, \sum_{j=1}^{n} \tilde{\lambda}_j Y_j^V, \tilde{Y}^R \right) \in T^{\mathrm{R}}_{\mathrm{VRS-TO}}.$$

Conditions (A.8a) and (A.8b) are a special case of equalities (A.7a) and (A.7b). By Lemma 3, the DMU $(\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R) \in T^{\mathrm{R}}_{\mathrm{VRS-TO}}$. □

Proof of Theorem 2. Consider any non-empty subset $J' \subseteq J = \{1, \ldots, n\}$. Define technology $T(J')$ as the set of all DMUs $(X, Y) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ for which there exists a vector $\lambda \in \mathbb{R}^n$, scalars π_t and ρ_{jl}, $\forall j, l, t$, and slack vectors S^V_Y, S^V_X, S^R_{Yj}, S^R_{Xj}, $\forall j$, such that

$$\sum_{j \in J'} \lambda_j Y^V_j + \sum_{t=1}^{K} \pi_t Q^V_t - S^V_Y = Y^V, \tag{A.9a}$$

$$\sum_{j \in J'} \lambda_j X^V_j + \sum_{t=1}^{K} \pi_t P^V_t + S^V_X = X^V, \tag{A.9b}$$

$$Y^R_j + \sum_{l=1}^{L} \rho_{jl} Q^R_l - S^R_{Yj} = Y^R, \qquad \forall j \in J', \tag{A.9c}$$

$$X^R_j + \sum_{l=1}^{L} \rho_{jl} P^R_l + S^R_{Xj} = X^R, \qquad \forall j \in J', \tag{A.9d}$$

$$\sum_{j \in J'} \lambda_j = 1, \tag{A.9e}$$

$$X^R \leq \bar{X}^R, \tag{A.9f}$$

$$Y^R \leq \bar{Y}^R, \tag{A.9g}$$

$$\lambda_j, \pi_t, \rho_{jl} \geq 0, \ \ \forall j \in J', \ \ t = 1, \ldots, K, \ \ l = 1, \ldots, L, \tag{A.9h}$$

$$S^V_Y, S^V_X, S^R_{Yj}, S^R_{Xj} \geq \mathbf{0}, \forall j \in J'. \tag{A.9i}$$

The proof now follows from Lemmas 5 and 6. □

Lemma 5. *Technology $T(J')$ defined by conditions (A.9) is a polyhedral set.*

Proof of Lemma 5. Restate conditions (A.9c) and (A.9d) by the following conditions:

$$\tilde{Y}^R = Y^R, \tag{A.10a}$$

$$\tilde{X}^R = X^R, \tag{A.10b}$$

$$Y^R_j + \sum_{l=1}^{L} \rho_{jl} Q^R_l - S^R_{Yj} - \tilde{Y}^R = \mathbf{0}, \qquad \forall j \in J', \tag{A.10c}$$

$$X^R_j + \sum_{l=1}^{L} \rho_{jl} P^R_l + S^R_{Xj} - \tilde{X}^R = \mathbf{0}, \qquad \forall j \in J', \tag{A.10d}$$

where $\tilde{X}^R \in \mathbb{R}^{m^R}$ and $\tilde{Y}^R \in \mathbb{R}^{s^R}$ are variable vectors. Consider the set C of all vectors whose components include variable vectors and scalars λ, π, ρ_{jl}, $j \in J'$, $l = 1, \ldots, L$, and $\tilde{X}^R$, $\tilde{Y}^R$, S^V_Y, S^V_X, S^R_{Yj}, S^R_{Xj}, $j \in J'$, that satisfy conditions (A.9e), (A.9h), (A.9i) and (A.10c), (A.10d). By definition, C is a polyhedral set (Rockafellar, 1970).

Define W as the set of all vectors $w = (Y^V, X^V, Y^R, X^R)$ obtainable by the linear transformation described by equalities (A.9a), (A.9b), (A.10a) and (A.10b) from the elements of C. By Theorem 19.3 in Rockafellar (1970), the set W is polyhedral. Then technology $T(J')$ is a polyhedral set as the intersection of the set W with the polyhedral set defined by the nonnegativity conditions $X^V, X^R, Y^V, Y^R \geq \mathbf{0}$ and conditions (A.9f) and (A.9g). □

Lemma 6. *Technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ is the union of the finite number of technologies $T(J')$ formed by all non-empty subsets $J' \subseteq J$.*

Proof of Lemma 6. Consider any DMU $(X, Y) \in T^{\mathrm{R}}_{\mathrm{VRS-TO}}$. By Theorem 1, it satisfies conditions (8) with some vector $\lambda' \in \mathbb{R}^n$ and scalars π'_t, $t = 1, \ldots, K$, and ρ'_{jl}, $j = 1, \ldots, n$, $l = 1, \ldots, L$. Define J' as the set of all $j \in J$ such that $\lambda'_j > 0$. Then (X, Y) satisfies conditions (A.9) with the vector λ obtained from λ' by omitting its zero components, and with the same scalars π'_t, and ρ'_{jl}, $\forall j, l, t$. Therefore, $(X, Y) \in T(J')$, and (X, Y) belongs to the union of all such technologies $T(J')$.

Conversely, let $(X, Y) \in T(J')$ for some set J'. Then (X, Y) satisfies (A.9) with some vector $\tilde{\lambda} \in \mathbb{R}^n$, scalars $\tilde{\pi}_t$ and $\tilde{\rho}_{jl}$, $\forall j, l, t$, and slack vectors $\tilde{S}^V_Y$, $\tilde{S}^V_X$, $\tilde{S}^R_{Yj}$, $\tilde{S}^R_{Xj}$, $\forall j \in J'$. Define vector $\lambda' \in \mathbb{R}^n$ as follows: $\lambda'_j = \tilde{\lambda}_j$, for all $j \in J'$, and $\lambda'_j = 0$, for all $j \in J \setminus J'$. Then (X, Y) satisfies conditions (8) with the so defined λ' and the same $\tilde{\pi}_t$ and $\tilde{\rho}_{jl}$, $\forall j, l, t$. Therefore, $(X, Y) \in T^{\mathrm{R}}_{\mathrm{VRS-TO}}$. □

Proof of Corollary 1. By Theorem 2, technology $T^{\mathrm{R}}_{\mathrm{VRS-TO}}$ is a finite union of polyhedral and, therefore, closed technologies. Their union is a closed set. □

Proof of Theorem 3. Denote T' the technology defined by conditions (10). More precisely, T' is the set of all DMUs $(X, Y) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ for which there exist vectors $\lambda, \alpha \in \mathbb{R}^n$ and scalars π_t, $t = 1, \ldots, K$, and ρ_{jl}, $j = 1, \ldots, n$, $l = 1, \ldots, L$, such that all conditions (10) are satisfied. We need to prove that $T^{\mathrm{F}}_{\mathrm{CRS-TO}} = T'$.

Let us first prove that T' satisfies Axioms 1–5. The proofs of Axioms 1,2 and 4 are similar to the case of R-VRS. The proofs that T' satisfies Axioms 3 and 5 are given separately, in Lemmas 7 and 8. Therefore, we have $T^{\mathrm{F}}_{\mathrm{CRS-TO}} \subseteq T'$. The opposite embedding $T' \subseteq T^{\mathrm{F}}_{\mathrm{CRS-TO}}$ is established by Lemma 9. The two proved embeddings imply that $T^{\mathrm{F}}_{\mathrm{CRS-TO}} = T'$. □

Lemma 7. *Technology T' satisfies Axiom 3.*

Proof of Lemma 7. Let DMUs $(\tilde{X}, \tilde{Y})$ and $(\hat{X}, \hat{Y})$ satisfy (10) with the combined vectors $(\tilde{\lambda}, \tilde{\alpha}, \tilde{\pi}; \tilde{\rho}_{jl} \mid \forall j, l)$ and $(\hat{\lambda}, \hat{\alpha}, \hat{\pi}; \hat{\rho}_{jl} \mid \forall j, l)$, respectively. Let the equalities (3) assumed by Axiom 3 be true. Select any $\gamma \in [0, 1]$ and define the convex combination

$$\left(\bar{X}^V, \bar{X}^R, \bar{Y}^V, \bar{Y}^R\right) = \gamma\left(\tilde{X}^V, \tilde{X}^R, \tilde{Y}^V, \tilde{Y}^R\right) + (1 - \gamma)\left(\hat{X}^V, \hat{X}^R, \hat{Y}^V, \hat{Y}^R\right).$$

Note that $\bar{X}$ and $\bar{Y}$ are nonnegative. It remains to be proved that DMU $(\bar{X}, \bar{Y})$ satisfies conditions (10) with some vectors $\bar{\lambda}$, $\bar{\alpha}$, $\bar{\pi}$ and scalars $\bar{\rho}_{jl}$.

First, as in the proof of Lemma 1, define vectors $\bar{\lambda}$ and $\bar{\pi}$ by formula (A.2). Further define the sets $\tilde{J}$ and $\hat{J}$ by formulae (A.3) and the multipliers $\bar{\rho}_{jl}$ by formula (A.4). To define vector $\bar{\alpha}$, first define the set $J^+ = \{j \in J \mid \bar{\lambda}_j > 0\} = \tilde{J} \cup \hat{J}$. Note that the set J^+ is not empty. For each $j \in J^+$, define $\bar{\alpha}_j$ from the following equality:

$$\bar{\lambda}_j \bar{\alpha}_j = \gamma \tilde{\lambda}_j \tilde{\alpha}_j + (1 - \gamma) \hat{\lambda}_j \hat{\alpha}_j. \tag{A.11}$$

Using (A.2) and (A.11), for each $j \in J^+$, we have

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

For each $j \in J \setminus J^+$ (i.e., for each j such that $\bar{\lambda}_j = 0$), we arbitrarily define $\bar{\alpha}_j = 1$. Let us prove that DMU $(\bar{X}, \bar{Y})$ satisfies (10) with the vectors $\bar{\lambda}$ $\bar{\alpha}$, $\bar{\pi}$ and scalars $\bar{\rho}_{jl}$.

To prove that inequalities (10a) and (10b) are true, we state them twice, for DMU $(\tilde{X}, \tilde{Y})$ and vectors $\tilde{\lambda}$, $\tilde{\alpha}$ and $\tilde{\pi}$, and for DMU $(\hat{X}, \hat{Y})$ and vectors $\hat{\lambda}$, $\hat{\alpha}$ and $\hat{\pi}$. Multiply these inequalities by γ and $1 - \gamma$, respectively. The proof is completed by adding the resulting inequalities and noting (A.2) and (A.11). The proof of the remaining conditions (10c)–(10h) is similar to their proof in Lemma 1. Therefore, $(\bar{X}, \bar{Y}) \in T'$. □

Lemma 8. *Technology T' satisfies Axiom 5.*

Proof of Lemma 8. Any DMU $(X^V, X^R, Y^V, Y^R) \in T'$ satisfies conditions (10) with some vectors λ, α and π, and scalars ρ_{jl}, $j = 1, \dots, n$, $l = 1, \dots, L$. Consider any scaling factor $\gamma \geq 0$ and define the scaled DMU $(\gamma X^V, X^R, \gamma Y^V, Y^R)$. This scaled DMU satisfies all conditions (10) with the vectors $\tilde{\lambda} = \lambda$, $\tilde{\alpha} = \gamma\alpha$ and the scalars $\tilde{\pi}_l = \gamma \pi_l$ and $\tilde{\rho}_{jt} = \rho_{jt}$, $\forall j, l, t$. It is, therefore, in T', and T' satisfies Axiom 5. □

Lemma 9. *The following embedding is true: $T' \subseteq T^{\mathrm{F}}_{\mathrm{CRS-TO}}$.*

Proof of Lemma 9. The proof follows closely the proof of Lemma 4, in which we replace the observed DMUs $\left(X^V_j, X^R_j, Y^V_j, Y^R_j\right)$ by their scaled analogues $(\alpha X^V_j, X^R_j, \alpha Y^V_j, Y^R_j)$. □

Proof of Theorem 4. The proof follows from Lemmas 10 and 11. □

Lemma 10. *Let DMU (X, Y) satisfy conditions (10) with vectors λ, α and scalars π_t and ρ_{jl}. Then (X, Y) satisfies conditions (13) with some vectors $\kappa, \mu, \nu \in \mathbb{R}^n$ and the same scalars π_t and ρ_{jl}.*

Proof of Lemma 10. Consider any $j \in J$ and let $\kappa_j = \lambda_j$. If $\lambda_j = 0$, let $\mu_j = \nu_j = 0$. If $\lambda_j > 0$, consider two cases. If $\alpha \geq 1$, define $\mu_j = \lambda_j\alpha_j - \lambda_j = \lambda_j(\alpha_j - 1)$ and $\nu_j = 0$. If $0 \leq \alpha < 1$, define $\mu_j = 0$ and $\nu_j = \lambda_j - \lambda_j\alpha_j = \lambda_j(1 - \alpha_j)$. In both cases, $\mu_j, \nu_j \geq 0$. Furthermore, we always have $\lambda_j\alpha_j = \kappa_j - \nu_j + \mu_j$ and $\kappa_j - \nu_j \geq 0$. Replacing $\lambda_j\alpha_j$ in inequalities (10a) and (10b) by the terms $\kappa_j - \nu_j + \mu_j$, for all $j \in J$, we observe that the DMU (X, Y) satisfies conditions (13) with the vectors κ, μ, ν and scalars π_t and ρ_{jl}. □

Lemma 11. *Let DMU (X, Y) satisfy conditions (13) with vectors κ, μ, ν and scalars π_t and ρ_{jl}. Then (X, Y) satisfies conditions (10) with some vectors $\lambda, \alpha \in \mathbb{R}^n$ and the same scalars π_t and ρ_{jl}.*

Proof of Lemma 11. Define $\Lambda = \sum_{j=1}^n (\kappa_j + \mu_j)$. By (13e), we have $\Lambda \geq 1$. For each $j \in J$, define $\lambda_j = (\kappa_j + \mu_j)/\Lambda$. Two further cases arise. If $\lambda_j > 0$, let $\alpha_j = (\kappa_j + \mu_j - \nu_j)/\lambda_j$. If $\lambda_j = 0$, then $\kappa_j = \mu_j = 0$ and, by (13f), $\nu_j = 0$. It does not matter how α_j is defined in this case. To be specific, let $\alpha_j = 1$. In both cases, we have $\lambda_j\alpha_j = \kappa_j - \nu_j + \mu_j$. It is straightforward to verify that conditions (13) stated for the vectors κ, μ, ν and scalars π_t and ρ_{jl}, imply conditions (10) stated in terms of vectors λ, α and scalars π_t and ρ_{jl}. □

Proof of Theorem 5. The proof is similar to the proof of Theorem 2 and requires a minor adjustment to Lemmas 5 and 6. In particular, we replace the statement (A.9) by a similar statement based on (13) and, in Lemma 6, we redefine the set J' as the set of all $j \in J$ such that $\kappa_j + \mu_j > 0$. The rest of the proof is similar and is omitted. □

Proof of Theorem 6. Let DMU $(X, Y) \in T^{\mathrm{F}}_{\mathrm{CRS-TO}}$. By Theorem 3, it satisfies conditions (10) with some vectors λ, α, π and scalars ρ_{jl}. Then (X, Y) satisfies all conditions (14) with the vector $\hat{\lambda}$ whose components are defined as $\hat{\lambda}_j = \lambda_j\alpha_j$, $\forall j \in J$, and the same vector π and scalars ρ_{jl}. In particular, $\hat{\lambda}_j > 0$ implies $\lambda_j > 0$, and conditions (14c) and (14d) follow from (10c) and (10d). Therefore, $(X, Y) \in \hat{T}^{\mathrm{F}}_{\mathrm{CRS-TO}}$.

Conversely, let $(X, Y) \in \hat{T}^{\mathrm{F}}_{\mathrm{CRS-TO}}$ satisfy conditions (14) with some vectors $\lambda \neq \mathbf{0}$, π and scalars ρ_{jl}. Then $\Lambda^* = \sum_{j \in J} \lambda_j > 0$. For all $j \in J$, define $\tilde{\lambda}_j = \lambda_j/\Lambda^*$ and $\tilde{\alpha}_j = \alpha_j\Lambda^*$. Then (X, Y) satisfies (10) with $\tilde{\lambda}$, $\tilde{\alpha}$ and the same π and ρ_{jl}. □

References

Allen, R., Athanassopoulos, A., Dyson, R. G., & Thanassoulis, E. (1997). Weights restrictions and value judgements in data envelopment analysis: Evolution, development and future directions. *Annals of Operations Research, 73*, 13–34.
Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*(9), 1078–1092.
Beale, E. M. L., & Tomlin, J. A. (1970). Special facilities in a general mathematical programming system for nonconvex problems using ordered sets of variables. In J. Lawrence (Ed.), *Proceedings of the fifth international conference on operational research* (pp. 447–454). London: Tavistock Publications.
Bradley, S., Johnes, G., & Millington, J. (2001). The effect of competition on the efficiency of secondary schools in England. *European Journal of Operational Research, 135*(3), 545–568.
Brennan, S., Haelermans, C., & Ruggiero, J. (2014). Nonparametric estimation of education productivity incorporating nondiscretionary inputs with an application to Dutch schools. *European Journal of Operational Research, 234*(3), 809–818.
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. (1970). *Data envelopment analysis. A comprehensive text with models, applications, references and DEA-Solver software (2nd ed.)*. New York: Springer.
Department for Education. (2022). Schools, pupils and their characteristics: Academic Year 2021/22. Department for Education.
Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C., & Shale, E. A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research, 132*(2), 245–259.
Emrouznejad, A., & Amin, G. R. (2009). DEA models for ratio data: Convexity consideration. *Applied Mathematical Modelling, 33*(1), 486–498.
Gorard, S. (2014). The link between academies in England, pupil outcomes and local patterns of socio-economic segregation between schools. *Research Papers in Education, 29*(3), 268–284.
Olesen, O. B., Petersen, N. C., & Podinovski, V. V. (2015). Efficiency analysis with ratio measures. *European Journal of Operational Research, 245*(2), 446–462.
Olesen, O. B., Petersen, N. C., & Podinovski, V. V. (2017). Efficiency measures and computational approaches for data envelopment analysis models with ratio inputs and outputs. *European Journal of Operational Research, 261*(2), 640–655.
Olesen, O. B., Petersen, N. C., & Podinovski, V. V. (2022a). Scale characteristics of variable returns-to-scale production technologies with ratio inputs and outputs. *Annals of Operations Research, 318*(1), 383–423.
Olesen, O. B., Petersen, N. C., & Podinovski, V. V. (2022b). The structure of production technologies with ratio inputs and outputs. *Journal of Productivity Analysis, 57*(3), 255–267.
Papaioannou, G., & Podinovski, V. V. (2023). Production technologies with ratio inputs and outputs. *European Journal of Operational Research, 310*(3), 1164–1178.
Podinovski, V. V. (2004). Production trade-offs and weight restrictions in data envelopment analysis. *Journal of the Operational Research Society, 55*(12), 1311–1322.
Podinovski, V. V. (2005). Selective convexity in DEA models. *European Journal of Operational Research, 161*(2), 552–563.
Podinovski, V. V., & Bouzdine-Chameeva, T. (2013). Weight restrictions and free production in data envelopment analysis. *Operations Research, 61*(2), 426–437.
Rockafellar, R. T. (1970). *Convex analysis*. Princeton, NJ: Princeton University Press.
Ross, A., Lessof, C., Brind, R., Khandker, R., & Aitken, D. (2020). *Examining the London advantage in attainment: Evidence from LSYPE*. Department for Education.
Silva, M. C. A., Camanho, A. S., & Barbosa, F. (2020). Benchmarking of secondary schools based on students' results in higher education. *Omega, 95*, 102119.
Thanassoulis, E., & Dunstan, P. (1994). Guiding schools to improved performance using data envelopment analysis: An illustration with data from a local education authority. *Journal of the Operational Research Society, 45*(11), 1247–1262.
Thanassoulis, E., Portela, M. C. S., & Despić, O. (2008). Data envelopment analysis: The mathematical programming approach to efficiency analysis. In H. O. Fried, C. A. K. Lovell, & S. S. Schmidt (Eds.), *The measurement of productive efficiency and productivity growth* (pp. 251–420). New York: Oxford University Press.