European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Decision support

# Production trade-offs in free disposal hull technologies

Mahmood Mehdiloo [a], Grammatoula Papaioannou [b], Victor V. Podinovski [b,*]

[a] Department of Mathematics and Applications, University of Mohaghegh Ardabili, Ardabil, Iran
[b] Loughborough Business School, Loughborough University, Leicestershire LE11 3TU, UK

ARTICLE INFO

*Keywords:*
Data envelopment analysis
Free disposal hull
Production trade-offs

ABSTRACT

In data envelopment analysis, production trade-offs are value judgements that represent simultaneous changes to the inputs and outputs assumed to be technologically possible for any production unit in the technology. The specification of production trade-offs generally leads to an enlargement of the model of technology and increasing its discriminating power on efficiency. In conventional convex variable and constant returns-to-scale models, production trade-offs are the dual forms of weight restrictions. In this paper, we extend the use of production trade-offs to the free disposal hull model of technology and its constant, non-increasing and non-decreasing returns-to-scale variants, in a single unifying development. We provide an axiomatic definition of the new nonconvex technologies, explore the notion of consistent trade-offs in such technologies and develop methods for its testing. We further develop different computational approaches for nonconvex models with production trade-offs. We illustrate the new models by an application in the context of higher education.

## 1. Introduction

In data envelopment analysis (DEA), the constant and variable returns-to-scale (CRS and VRS) models of Charnes, Cooper, and Rhodes (1978) and Banker, Charnes, and Cooper (1984) can be stated in the two mutually dual forms, known as the multiplier and envelopment models. Both forms allow the incorporation of value judgements as a way to refine the model (by the specification of additional information relevant to the efficiency assessment) and improve its discriminating power.

### 1.1. Production trade-offs in convex DEA models

In the multiplier model, value judgements are stated by weight restrictions—see, e.g., Allen, Athanassopoulos, Dyson, and Thanassoulis (1997), Thanassoulis, Portela, and Despić (2008), and Ramón, Ruiz, and Sirvent (2016). They typically represent the managerial perception of the relative importance of inputs and outputs in the assessment of efficiency of decision making units (DMUs). Production trade-offs are the dual forms of weight restrictions that appear in the envelopment models (Podinovski, 2004d). Such trade-offs are interpretable as simultaneous changes to the inputs and outputs that are assumed technologically possible for all DMUs in the technology.

Each production trade-off specifies technologically possible simultaneous changes for two and more inputs or outputs. For example, in the application to universities discussed in Section 7, we assume that the teaching of one undergraduate student at any department does not incur a higher cost to the university than the teaching one postgraduate student at the same department. This creates a simple trade-off stating that, for any university, it is technologically possible to increase the number of undergraduate (UG) students by one and simultaneously reduce the number of postgraduate (PG) students by one, without seeking extra resources and without any detriment to the other outputs. This trade-off does not mean that, in reality, any department can implement this change but rather that the reduction of one PG student releases sufficient costs that can be used to teach an extra UG student, should this be required.

Production trade-offs are related to various marginal characteristics of production frontiers, such as the rates of transformation, substitution and marginal productivity involving different measures (inputs and outputs). However, while such marginal rates are generally different at different efficient DMUs, the way production trade-offs are incorporated in the envelopment model means that they need to be acceptable for all DMUs in the technology. In this sense, production trade-offs should be *safe assumptions* that are intentionally less demanding than *any* of the precise (and often unknown) marginal rates evaluated at different points of the efficient frontier. For example, in the already mentioned university context, we may find out that the exact rate of substitution equates the costs of teaching one PG student to anything between 1 and 3 UG students, depending on the university or department. Then the trade-off stating that the reduction of one PG

* Corresponding author.
*E-mail addresses:* m.mehdiloo@uma.ac.ir (M. Mehdiloo), g.papaioannou@lboro.ac.uk (G. Papaioannou), v.podinovski@lboro.ac.uk (V.V. Podinovski).

https://doi.org/10.1016/j.ejor.2025.06.032
Received 16 December 2024; Accepted 25 June 2025
Available online 7 July 2025


Please cite this article as: Mahmood Mehdiloo et al., *European Journal of Operational Research,* https://doi.org/10.1016/j.ejor.2025.06.032

student releases sufficient resources to teach one UG student should be applicable to all universities.

The incorporation of production trade-offs in the model leads to an expansion of the underlying production technology in which any DMUs is regarded producible and can be explained based on the stated assumptions, including the specified trade-offs. In particular, we can explain why the radial or efficient target of any inefficient DMU should be considered technologically feasible—see, e.g., Atici and Podinovski (2015) for an example of such explanation in an application.

### 1.2. Free disposal hull and its variants

Free disposal hull (FDH) was introduced by Deprins, Simar, and Tulkens (1984) as a model of technology which incorporates the observed DMUs and is based on the single assumption of strong (free) disposability of all inputs and outputs. The FDH model was subsequently discussed by Tulkens (1993) and Tulkens and Vanden Eeckaut (1995). Its further developments were pursued in several directions. Cherchye, Kuosmanen, and Post (2000), Kerstens and Van de Woestyne (2021) and Kerstens, Sadeghi, Toloo, and Van de Woestyne (2022) explore the economic meaning of FDH and the cost function defined by this model. Tavakoli and Mostafaee (2019) consider a two-stage FDH, while Papaioannou and Podinovski (2024) introduce an FDH model with multiple component processes.

Following an outline of Bogetoft (1996), Kerstens and Vanden Eeckaut (1999) develop further variants of the FDH technology under the assumptions of constant, nonincreasing and nondecreasing (CRS, NIRS and NDRS) returns to scale. In line with the terminology used in the conventional convex case, the FDH technology and its variants under different assumptions of returns to scale are also referred to as the nonconvex VRS, CRS, NIRS and NDRS technologies. The last three of them are useful for the assessment of scale efficiency and most productive scale size (MPSS) of the DMUs in the FDH (nonconvex VRS) technology, which is conceptually similar to their evaluation in the case of convex VRS model (Banker, 1984). Podinovski (2004a, 2004b) shows that they are also useful for the characterization of global returns to scale whose types are indicative of the direction of resizing that a DMU should undertake on its way to MPSS.

The FDH model benchmarks any DMU against its actual observed peers, and not against hypothetical DMUs obtained as convex combinations of the observed DMUs which are usually explained by the (time) divisibility principle. The latter becomes problematic if production requires a significant start-up investment or exhibits other types of indivisibilities (Scarf, 1981). An example of this in the context of multi-stage production was discussed by Tone and Sahoo (2003). Furthermore, as noted by Koopmans (1957), assuming that inactivity (i.e., producing zero output from a zero input) is possible and that the technology is convex makes the increasing returns-to-scale characterization of efficient frontiers impossible. This contradicts the fact that in reality larger firms often exhibit higher productivity than smaller ones. Resolving this difficulty can be achieved by removing the assumption that the technology is convex, e.g., by using the FDH model of technology instead of its convex VRS analogue.

Although the above arguments make the FDH model attractive in applications, this comes at the expense of its relatively low discriminating power, compared to the convex VRS model. We may, however, consider the efficiency scores obtained from the FDH model as the upper bounds on the “true” efficiency of the DMUs. If a DMU is inefficient in the FDH model, then its efficiency cannot be improved in any other model that assumes free disposability of inputs and outputs.

### 1.3. Contribution

In this paper we consider an extension of the FDH model of technology by the specification of production trade-offs. In a single unifying development, we also simultaneously obtain its CRS, NIRS and NDRS variants. As mentioned above, these variants are useful for the analysis of scale characteristics and returns to scale of the FDH technology with trade-offs.

We make several contributions of theoretical and practical importance. First, we provide a full axiomatic derivation of the new nonconvex technologies which firmly places them in the realm of production theory. Second, we obtain three different statements of these technologies, each of which is useful for different theoretical and computational purposes. For example, this allows us to investigate properties of the new nonconvex technologies. In particular, it turns out that, although the convex VRS technology is the convex hull of FDH, this is no longer true for the same technologies extended by production trade-offs.

Third, we explore the notion of consistent production trade-offs and methods of its testing. Podinovski and Bouzdine-Chameeva (2013, 2015) define consistent trade-offs and dual weight restrictions as those that do not generate free or unlimited production in the technology. If trade-offs are inconsistent, this means that we made an error in their specification and need to reassess our value judgements. In our paper, we prove that, if the specified trade-offs are consistent in any one of the four nonconvex (VRS, CRS, NIRS or NDRS) technologies, then they are consistent in the other three of these four technologies. We also develop exact and sufficient computational tests of consistency of trade-offs.

Fourth, we develop two different computational approaches to solving (generally mixed integer nonlinear) models based on the four nonconvex (VRS, CRS, NIRS or NDRS) technologies with production trade-offs. One of these methods requires solving a single linear program for each DMU under the assessment.

Finally, we provide an illustrative application in the context of higher education that demonstrates an increasing discriminating power arising from the specification of production trade-offs in nonconvex models.

We proceed as follows. In Section 2, we give an overview of production trade-offs in convex DEA technologies. In Section 3, we define nonconvex technologies with production trade-offs and develop their alternative statements. In Section 4, we develop the notion of consistent trade-offs in nonconvex technologies and tests of consistency. In Sections 5 and 6, we consider computational approaches for the nonconvex models. In Section 7, we illustrate the increased discriminating power of the new models with trade-offs in the context of higher education. In Section 8, we provide concluding remarks.

Appendices A–C include additional discussion and proofs of mathematical results.

## 2. Convex technologies with production trade-offs

In this section, we give a unifying axiomatic definition of the four conventional convex DEA technologies and their extensions by production trade-offs. The terminology and notation used in this approach are related to the terminology used by Briec, Kerstens, and Vanden Eeckaut (2004). This allows us to streamline and simplify the development of analogous nonconvex technologies with production trade-offs in the subsequent sections.

### 2.1. Basic notation and terminology

Consider a production technology $\mathcal{T} \subset \mathbb{R}^m_+ \times \mathbb{R}^s_+$ with $m \geq 1$ inputs and $s \geq 1$ outputs. It consists of DMUs $(\mathbf{x}, \mathbf{y})$, where $\mathbf{x} \in \mathbb{R}^m_+$ and $\mathbf{y} \in \mathbb{R}^s_+$ are the vectors of inputs and outputs, respectively. Conceptually, technology $\mathcal{T}$ is interpreted as follows:

$$\mathcal{T} = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+ \mid \mathbf{x} \text{ can produce } \mathbf{y} \right\}.$$

Throughout this paper, we consider a finite set of observed DMUs, denoted $(\mathbf{x}_j, \mathbf{y}_j)$, $j \in \mathcal{J} = \{1, \ldots, J\}$. We denote $(\mathbf{x}_o, \mathbf{y}_o)$ the DMU under the evaluation. We assume that $\mathbf{x}_j \neq 0$ and $\mathbf{y}_j \neq 0$, for all $j \in \mathcal{J} \cup \{o\}$, i.e., each observed DMU and the DMU $(\mathbf{x}_o, \mathbf{y}_o)$ have at least one strictly positive input and at least one strictly positive output.

The vectors denoted $\mathbf{0}$ and $\mathbf{1}$ have all their components equal to 0 and 1, respectively. The dimensions of such vectors are defined by the context in which they are used. Inequalities stated for vectors are understood to be true for each of their components. For example, if $\hat{\mathbf{x}}$ and $\tilde{\mathbf{x}}$ are two vectors of inputs, then the vector inequality $\hat{\mathbf{x}} \geq \tilde{\mathbf{x}}$ means that $\hat{x}_i \geq \tilde{x}_i$, for all their components $i = 1, \ldots, m$.

### 2.2. Unified definitions

The CRS, NIRS and NDRS technologies are extensions of the VRS technology of Banker et al. (1984) defined by different assumptions about the scalability of DMUs. For all four technologies, we use a unifying notation $\mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS}}$, where $\Delta \in \{\mathrm{NI}, \mathrm{V}, \mathrm{ND}, \mathrm{C}\}$. For example, if $\Delta = \mathrm{NI}$, we have the NIRS technology $\mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS}} = \mathcal{T}^{\mathrm{C}}_{\mathrm{NIRS}}$.

Following the minimum extrapolation principle of Banker et al. (1984), technology $\mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS}}$ is defined as the intersection of all technologies $\mathcal{T} \subset \mathbb{R}^m_+ \times \mathbb{R}^s_+$ that satisfy the following four axioms. (We use the superscript C to identify a convex technology.)

**Axiom IO** (*Inclusion of Observations*) $(\mathbf{x}_j, \mathbf{y}_j) \in \mathcal{T}$ for all $j \in \mathcal{J}$.

**Axiom SD** (*Strong Disposability*) If $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}$, then $(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \in \mathcal{T}$ for all $\hat{\mathbf{x}} \geq \mathbf{x}$ and all $\mathbf{0} \leq \hat{\mathbf{y}} \leq \mathbf{y}$.

**Axiom ΔRS** (*Δ Returns to Scale*) $\delta\mathcal{T} \subseteq \mathcal{T}$ for all $\delta \in \mathcal{I}_\Delta$, where $\Delta \in \{\mathrm{NI}, \mathrm{V}, \mathrm{ND}, \mathrm{C}\}$ and $\mathcal{I}_{\mathrm{NI}} = [0, 1]$, $\mathcal{I}_{\mathrm{V}} = \{1\}$, $\mathcal{I}_{\mathrm{ND}} = [1, +\infty)$, $\mathcal{I}_{\mathrm{C}} = \mathbb{R}_+$.

**Axiom CT** (*Convexity of Technology*) $\mathcal{T}$ is convex.

We can state technology $\mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS}}$ in its conventional operational form as follows:

$$\mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS}} = \Big\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+ \mid \exists \lambda \in \mathbb{R}^J : \sum_{j \in \mathcal{J}} \lambda_j \mathbf{x}_j \leq \mathbf{x},\ \sum_{j \in \mathcal{J}} \lambda_j \mathbf{y}_j \geq \mathbf{y},\ \mathbf{1}^\top \lambda \in \mathcal{I}_\Delta,\ \lambda \geq \mathbf{0} \Big\}, \tag{1}$$

where the notation $\mathcal{I}_\Delta$ is as introduced in the statement of Axiom ΔRS. For example, if $\Delta = \mathrm{C}$, then $\mathcal{I}_{\mathrm{C}} = \mathbb{R}_+$ and the condition $\mathbf{1}^\top \lambda \in \mathcal{I}_\Delta$ becomes redundant and can be removed. This defines the standard CRS technology. If $\Delta = \mathrm{NI}$, then $\mathcal{I}_{\mathrm{NI}} = [0, 1]$ and the condition $\mathbf{1}^\top \lambda \in \mathcal{I}_\Delta$ becomes $0 \leq \mathbf{1}^\top \lambda \leq 1$, which defines the NIRS technology.

### 2.3. Production trade-offs in convex technologies

Following Podinovski (2004d), a production trade-off is a statement of simultaneous changes to the inputs and outputs that are assumed technologically possible at any DMU in technology $\mathcal{T}$ (provided the inputs and outputs of resulting DMU remain nonnegative). Suppose we have specified $K$ production trade-offs. Formally, these can be stated as $K$ pairs of vectors

$$(\mathbf{p}_k, \mathbf{q}_k), \quad k \in \mathcal{K} = \{1, \ldots, K\}, \tag{2}$$

where the vectors $\mathbf{p}_k \in \mathbb{R}^m$ and $\mathbf{q}_k \in \mathbb{R}^s$ represent simultaneous changes to the inputs and outputs of the DMUs, respectively.

The application of any single trade-off to any DMU $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}$ creates another DMU in the technology, to which we can further apply the same or any other of the trade-offs. Allowing a fractional number of times $\pi_k \geq 0$ that each trade-off (2) is applied, we state the resulting DMU as

$$(\mathbf{x}', \mathbf{y}') = (\mathbf{x}, \mathbf{y}) + \sum_{k \in \mathcal{K}} \pi_k (\mathbf{p}_k, \mathbf{q}_k). \tag{3}$$

The assumption that production trade-offs (2) represent feasible simultaneous changes to the inputs and outputs of any DMU in the technology is stated by the following axiom.

**Axiom FTO** (*Feasibility of Trade-Offs*) If DMU $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}$, then the DMU $(\mathbf{x}', \mathbf{y}')$ defined by (3) is in $\mathcal{T}$, provided $\mathbf{x} + \sum_{k \in \mathcal{K}} \pi_k \mathbf{p}_k \geq \mathbf{0}$ and $\mathbf{y} + \sum_{k \in \mathcal{K}} \pi_k \mathbf{q}_k \geq \mathbf{0}$.

**Remark 1.** Podinovski (2004d) uses a different variant of Axiom FTO which assumes that a modification of DMU $(\mathbf{x}, \mathbf{y})$ by any *single* trade-off (2) is technologically possible. In Appendix A, we compare the two axioms and show that Axiom FTO is generally a stronger assumption than the original axiom of Podinovski (2004d). However, as follows from the proof of Theorem 1 in Podinovski (2004d), in any closed technology, the two axioms are equivalent. Let us also note that the above Axiom FTO is a special case of the axiom of production trade-offs used by Podinovski, Wu, and Argyris (2024) for the technology with both volume and ratio types of inputs and outputs.

The specification of production trade-offs (2) for technology $\mathcal{T}$ generates additional (unobserved) DMUs that are assumed to be technologically possible. Using the minimum extrapolation principle, we define technology $\mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS\text{-}TO}}$, which extends the technology $\mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS}}$ by the application of trade-offs (2), as the intersection of all technologies $\mathcal{T} \subset \mathbb{R}^m_+ \times \mathbb{R}^s_+$ that satisfy Axioms IO, SD, CT, ΔRS and FTO.

Podinovski (2004d) establishes operational statements of technology $\mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS\text{-}TO}}$ in the cases of CRS and VRS. It is straightforward to verify that the mathematical derivation of these statements extends to the cases of NIRS and NDRS, with obvious minor modifications. This results in the following unified statement of technology $\mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS\text{-}TO}}$, which includes the cases of VRS, CRS, NIRS and NDRS:

$$\mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS\text{-}TO}} = \Big\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+ \mid \exists \lambda \in \mathbb{R}^J,\ \boldsymbol{\pi} \in \mathbb{R}^K : \sum_{j \in \mathcal{J}} \lambda_j \mathbf{x}_j + \sum_{k \in \mathcal{K}} \pi_k \mathbf{p}_k \leq \mathbf{x},\ \sum_{j \in \mathcal{J}} \lambda_j \mathbf{y}_j + \sum_{k \in \mathcal{K}} \pi_k \mathbf{q}_k \geq \mathbf{y},\ \mathbf{1}^\top \lambda \in \mathcal{I}_\Delta,\ \lambda \geq \mathbf{0},\ \boldsymbol{\pi} \geq \mathbf{0} \Big\}. \tag{4}$$

We conclude this section by stating the embedding

$$\mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS}} \subseteq \mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS\text{-}TO}}. \tag{5}$$

This embedding is easy to prove. Indeed, any DMU $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS}}$ satisfies conditions of the statement (1) with some vector $\lambda'$. Then it also satisfies all conditions of the statement (4) of technology $\mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS\text{-}TO}}$ with the same vector $\lambda'$ and vector $\boldsymbol{\pi} = \mathbf{0}$. Therefore, $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}^{\mathrm{C}}_{\Delta\mathrm{RS\text{-}TO}}$, and the embedding (5) follows.

## 3. Nonconvex technologies with production trade-offs

In this section, we develop an extension of the notion of production trade-offs to the FDH model of technology and its variants for the cases of CRS, NIRS and NDRS.

### 3.1. Axiomatic definitions

Removing the assumption of convexity and using the minimum extrapolation principle, we define the nonconvex technology $\mathcal{T}^{\mathrm{NC}}_{\Delta\mathrm{RS}}$ as the intersection of all technologies $\mathcal{T} \subset \mathbb{R}^m_+ \times \mathbb{R}^s_+$ that satisfy Axioms IO, SD and ΔRS. (The superscript NC identifies the nonconvex case).

A conventional statement (see, e.g., Briec et al., 2004) of this technology is

$$\mathcal{T}^{\mathrm{NC}}_{\Delta\mathrm{RS}} = \Big\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+ \mid \exists \lambda \in \mathbb{R}^J, \delta \in \mathbb{R} : \sum_{j \in \mathcal{J}} \delta \lambda_j \mathbf{x}_j \leq \mathbf{x},\ \sum_{j \in \mathcal{J}} \delta \lambda_j \mathbf{y}_j \geq \mathbf{y},\ \mathbf{1}^\top \lambda = 1,\ \lambda \in \{0, 1\}^J, \delta \in \mathcal{I}_\Delta \Big\}. \tag{6}$$

For example, if $\Delta = \text{V}$, then $\delta = 1$ and technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS}}$ is the standard FDH technology. If $\Delta = \text{C}$, then $\delta \in \mathbb{R}_+$, and we obtain the CRS variant of the FDH technology.

Let us now consider an extension of the nonconvex technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS}}$ by the incorporation of production trade-offs (2). Using the minimum extrapolation principle, we define the extended technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ as the intersection of all technologies $\mathcal{T} \subset \mathbb{R}^m_+ \times \mathbb{R}^s_+$ that satisfy Axioms IO, SD, ΔRS and FTO.

Because technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ satisfies Axioms IO, SD and ΔRS, and because technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS}}$ is the smallest technology satisfying the same axioms, we have the following embedding:

$$\mathcal{T}^{\text{NC}}_{\Delta\text{RS}} \subseteq \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}. \tag{7}$$

The above axiomatic definition of technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ is not operational and cannot be used in computations. Below we develop three different approaches that operationalize this definition. Each of these approaches is useful for different tasks.

**Remark 2.** Technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ includes the cases of the standard FDH technology extended by production trade-offs (2) and its three (CRS, NIRS and NDRS) variants. The extended FDH technology may arguably be the most interesting for practical applications in which we are primarily interested in the assessment of efficiency of the DMUs. The CRS, NIRS and NDRS technologies are important for the evaluation of MPSS of a DMU, as introduced to DEA by Banker (1984), and the direction of resizing that a DMU should undertake in its movement towards MPSS.

For *convex* technologies (which do not include technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$), such approach was developed by Färe, Grosskopf, and Lovell (1983, 1985) as a qualitative method of evaluation of local returns to scale (RTS), which is consistent with the notion of scale elasticity but does not require its calculation. (The local characterization of RTS based on the calculation of one-sided scale elasticities was considered in detail by Banker and Thrall (1992), Hadjicostas and Soteriou (2006), Førsund, Hjalmarsson, Krivonozhko, and Utkin (2007), Sahoo and Tone (2015), and Podinovski, Chambers, Atici, and Deineko (2016).)

Podinovski (2004a, 2004b) modifies this approach for the case of arbitrary *nonconvex* technologies, which applies to all technologies $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ and leads to the notion of global returns to scale as an indicator of the direction to MPSS. For nonconvex technologies, the local and global RTS are generally two different characterizations of the production frontier. We illustrate the use of the nonconvex CRS, NIRS and NDRS technologies for the evaluation of global RTS in technology $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$ in the application in Section 7.

### 3.2. A basic statement

We can obtain the statement of technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ by a simple adaptation of the statement (6) of technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS}}$ in which we incorporate production trade-offs (2) as in the statement (4) of its convex analogue technology $\mathcal{T}^{\text{C}}_{\Delta\text{RS-TO}}$.

**Theorem 3.1.** *Technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ is equivalently stated as follows:*

$$\begin{aligned} \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}} = \Big\{ & (\mathbf{x},\mathbf{y}) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+ \mid \\ & \exists \lambda \in \mathbb{R}^J, \delta \in \mathbb{R}, \boldsymbol{\pi} \in \mathbb{R}^K : \\ & \sum_{j\in\mathcal{J}} \delta\lambda_j \mathbf{x}_j + \sum_{k\in\mathcal{K}} \pi_k \mathbf{p}_k \le \mathbf{x}, \\ & \sum_{j\in\mathcal{J}} \delta\lambda_j \mathbf{y}_j + \sum_{k\in\mathcal{K}} \pi_k \mathbf{q}_k \ge \mathbf{y}, \\ & \mathbf{1}^\top \lambda = 1,\ \lambda \in \{0,1\}^J, \delta \in \mathcal{I}_\Delta, \boldsymbol{\pi} \ge \mathbf{0} \Big\}. \end{aligned} \tag{8}$$

The above statement of technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ is in line with traditional statements of DEA technologies and is therefore intuitive and easy for interpretation. In the case of VRS, the condition $\delta \in \mathcal{I}_\Delta$ becomes $\delta = 1$, and both input and output inequalities become linear. Assessing the input or output radial efficiency of a DMU using this statement of technology $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$ requires solving a mixed integer linear program.

In the cases of CRS, NIRS and NDRS, the statement (8) includes nonlinear terms $\delta\lambda_j$ in the input and output inequalities and becomes problematic from the computational perspective. One way to linearize such statement is by the use of the substitution of variables and "Big M" approach, as developed by Podinovski (2004c) for the statement (6) of technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS}}$ (without trade-offs).

### 3.3. A decomposition-based statement

Briec et al. (2004) note that, in the conventional case without trade-offs, technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS}}$ is the union of $J$ elementary "subtechnologies" $\mathcal{T}^{\text{NC}}_{\Delta\text{RS}}(\mathbf{x}_j,\mathbf{y}_j)$ each of which is generated by a single observed DMU $(\mathbf{x}_j,\mathbf{y}_j)$, $j \in \mathcal{J}$. Below we show that this approach extends to the case involving production trade-offs (2). The statement of each such elementary subtechnology is given by linear inequalities, which is attractive from the computational point of view.

To state this formally, for each $j \in \mathcal{J}$, we denote $\mathcal{H}^j_{\Delta\text{RS-TO}}$ the Δ-strong disposal hull of the observed DMU $(\mathbf{x}_j,\mathbf{y}_j)$ that incorporates production trade-offs (2):

$$\begin{aligned} \mathcal{H}^j_{\Delta\text{RS-TO}} = \Big\{ & (\mathbf{x},\mathbf{y}) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+ \mid \exists\, \delta \in \mathbb{R}, \boldsymbol{\pi} \in \mathbb{R}^K : \\ & \delta\mathbf{x}_j + \sum_{k\in\mathcal{K}} \pi_k \mathbf{p}_k \le \mathbf{x}, \\ & \delta\mathbf{y}_j + \sum_{k\in\mathcal{K}} \pi_k \mathbf{q}_k \ge \mathbf{y}, \\ & \delta \in \mathcal{I}_\Delta, \boldsymbol{\pi} \ge \mathbf{0} \Big\}. \end{aligned} \tag{9}$$

We now have the following representation of technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ which extends a similar representation of technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS}}$ obtained by Briec et al. (2004) for the standard case without production trade-offs.

**Theorem 3.2.** *Technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ is equivalently stated as follows:*

$$\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}} = \bigcup_{j\in\mathcal{J}} \mathcal{H}^j_{\Delta\text{RS-TO}}. \tag{10}$$

**Remark 3.** Theorem 3.2 restates technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ as a finite union of the hulls $\mathcal{H}^j_{\Delta\text{RS-TO}}$, $j \in \mathcal{J}$. Depending on Δ, each of these hulls can be regarded as the convex VRS, NIRS, NDRS or CRS technology generated by the single observed DMU $(\mathbf{x}_j,\mathbf{y}_j)$, $j \in \mathcal{J}$, and the trade-offs (2). This means that the trade-offs (2) allow the standard interpretation as restrictions on the normal vectors of the supporting hyperplanes to the individual hulls, or as weight restrictions in the corresponding multiplier models based on them (Podinovski, 2004d).

The statement (10) of technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ is useful for its graphical depictions as the union of the hulls $\mathcal{H}^j_{\Delta\text{RS-TO}}$ defined by (9), each of which is straightforward to visualize. We illustrate this by the following example.

**Example 1.** Let $A = (2,2)$, $B = (6,6)$ and $C = (8,4)$ be three observed DMUs, where the first component is input $x$ and the second is output $y$. Also consider the following two trade-offs:

$$(p_1,q_1) = (-1,-2), \quad (p_2,q_2) = (4,2). \tag{11}$$

The nonconvex technology $\mathcal{T}^{\text{NC}}_{\text{VRS}}$ generated by the observed DMUs $A$, $B$ and $C$ is shown in Fig. 1 as the darker shaded area below and to the right of the line $UAFBV$.

The areas below and to the right of the lines $A'AA''$, $B'BB''$ and $C'CC''$ show the three hulls $\mathcal{H}^A_{\text{VRS-TO}}$, $\mathcal{H}^B_{\text{VRS-TO}}$ and $\mathcal{H}^C_{\text{VRS-TO}}$, respectively. For example, the point $C^*$ is obtained by the application of trade-off $(p_1,q_1)$ to DMU $B$ in proportion $\pi_1 = 1$ and is therefore in the hull $\mathcal{H}^B_{\text{VRS-TO}}$.

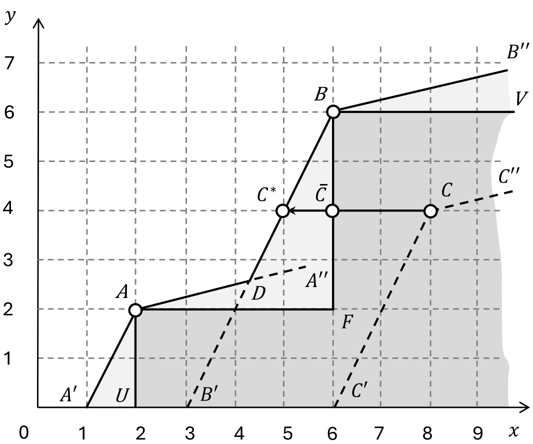


**Fig. 1.** Technology $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$ in Example 1.

By Theorem 3.2, technology $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$ is the union of these three hulls, i.e., the total shaded area below the line $A'ADBB''$.

It is worth highlighting that, in line with the embedding (7), the specification of trade-offs (11) has expanded the underlying technology $\mathcal{T}^{\text{NC}}_{\text{VRS}}$. As a result, for example, the input radial target of DMU $C$ moves from the point $\bar{C}$ in technology $\mathcal{T}^{\text{NC}}_{\text{VRS}}$ to point $C^*$ in technology $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$, and the input radial efficiency of DMU $C$ decreases from $6/8 = 0.75$ to $5/8 = 0.625$.

In Sections 5 and 6, we show that the statement (10) is also useful for computational purposes, as it requires solving $J$ linear programs for the assessment of efficiency of each DMU.

The statement (10) can be used for establishing some properties of technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$, which are more difficult to obtain from its alternative statements. As an example, let us prove that technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ is a closed set.

**Theorem 3.3.** *For each $j \in \mathcal{J}$, the hull $\mathcal{H}^{j}_{\Delta\text{RS-TO}}$ is a polyhedral and, therefore, closed and convex set.*

**Corollary 3.1.** *Technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ is a closed set.*

It is a simple mathematical fact that the conventional convex VRS technology is the convex hull of the FDH technology, i.e., $\text{conv}\left(\mathcal{T}^{\text{NC}}_{\text{VRS}}\right) = \mathcal{T}^{\text{C}}_{\text{VRS}}$. Interestingly, the same equality is generally not true for technologies incorporating trade-offs (2) and we only have the embedding (which follows from the embedding $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}} \subseteq \mathcal{T}^{\text{C}}_{\text{VRS-TO}}$):

$$\text{conv}\left(\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}\right) \subseteq \mathcal{T}^{\text{C}}_{\text{VRS-TO}}. \tag{12}$$

The following is an example with two inputs and two outputs in which the embedding (12) is proper.

**Example 2.** Consider the two observed DMUs $(\mathbf{x}_1, \mathbf{y}_1) = (2, 0, 10, 1)^\top$ and $(\mathbf{x}_2, \mathbf{y}_2) = (0, 2, 10, 1)^\top$, and the single trade-off $(\mathbf{p}, \mathbf{q}) = (-0.5, -0.5, -5, -1)^\top$, where the first two components are inputs and the last two components are outputs. Consider the convex and nonconvex VRS technologies $\mathcal{T}^{\text{C}}_{\text{VRS-TO}}$ and $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$ generated by the above observed DMUs and the trade-off, and the convex hull $\text{conv}\left(\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}\right)$ of the nonconvex technology.

The simple average of DMUs $(\mathbf{x}_1, \mathbf{y}_1)$ and $(\mathbf{x}_2, \mathbf{y}_2)$ is the DMU $(1, 1, 10, 1)^\top$. Modifying this DMU by the trade-off $(\mathbf{p}, \mathbf{q})$, we obtain the DMU $(\mathbf{x}', \mathbf{y}') = (0.5, 0.5, 5, 0)^\top \in \mathcal{T}^{\text{C}}_{\text{VRS-TO}}$.

Let us show that DMU $(\mathbf{x}', \mathbf{y}')$ is not in the convex hull $\text{conv}\left(\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}\right)$. Indeed, any DMU $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$ is located in either hull $\mathcal{H}^{j}_{\text{VRS-TO}}$, $j = 1, 2$, or in both of them. Suppose that $(\mathbf{x}, \mathbf{y}) \in \mathcal{H}^{1}_{\text{VRS-TO}}$. According to (9) in which $\delta = 1$, and taking into account that $\mathbf{y} \geq \mathbf{0}$, observe that the multiplier $\pi$ used with the trade-off $(\mathbf{p}, \mathbf{q})$ cannot be larger than 1 (as otherwise the total on the left-hand side of the inequality $1 + \pi(-1) \geq y_2$ for output 2 becomes negative). Then input 1 of DMU $(\mathbf{x}, \mathbf{y})$ is greater than or equal to $2 - 0.5 = 1.5$. Similarly, input 2 of any DMU in the hull $\mathcal{H}^{2}_{\text{VRS-TO}}$ is greater than or equal to 1.5. Therefore, the sum of inputs 1 and 2 of any DMU in technology $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}} = \mathcal{H}^{1}_{\text{VRS-TO}} \cup \mathcal{H}^{2}_{\text{VRS-TO}}$ is greater than or equal to 1.5. The same is true for any DMU in the convex hull $\text{conv}\left(\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}\right)$. This means that the DMU $(\mathbf{x}', \mathbf{y}') = (0.5, 0.5, 5, 0)^\top$, for which the sum of two inputs is equal to 1, is not in $\text{conv}\left(\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}\right)$.

It turns out that, in the case of a single input $x$ and a single output $y$, the embedding (12) becomes an equality. We prove this result under the assumption that technology $\mathcal{T}^{\text{C}}_{\text{VRS-TO}}$ does not allow unlimited production of output $y$ from any fixed input $x$, which is a standard production assumption (see, e.g., Färe et al., 1985). (We discuss this assumption further in Section 4.)

**Theorem 3.4.** *In the case of a single input $x$ and a single output $y$, the embedding (12) is satisfied as an equality, i.e., $\text{conv}\left(\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}\right) = \mathcal{T}^{\text{C}}_{\text{VRS-TO}}$.*

### 3.4. A linearized statement

Technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ can also be represented using a single linearized statement, using the idea of Afsharian and Podinovski (2018) and its extension to the whole class of technologies $\mathcal{T}^{\text{NC}}_{\Delta\text{RS}}$ by Mehdiloo, Sadeghi, and Kerstens (2024). From a computational perspective, using this statement for the efficiency assessment requires solving a single linear program for each DMU under the assessment, which offers clear computational advantages.

**Theorem 3.5.** *Technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ is equivalently stated as follows:*

$$\begin{aligned}
\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}} = \Big\{ & (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+ \mid \\
& \exists \boldsymbol{\lambda}, \boldsymbol{\eta} \in \mathbb{R}^J, \boldsymbol{\pi}_j \in \mathbb{R}^K, j \in \mathcal{J}: \\
& \lambda_j \mathbf{x}_j + \sum_{k \in \mathcal{K}} \pi_{jk} \mathbf{p}_k \leq \eta_j \mathbf{x}, \ j \in \mathcal{J}, \\
& \lambda_j \mathbf{y}_j + \sum_{k \in \mathcal{K}} \pi_{jk} \mathbf{q}_k \geq \eta_j \mathbf{y}, \ j \in \mathcal{J}, \\
& \lambda_j - \eta_j \in \mathcal{R}_\Delta, \ j \in \mathcal{J}, \\
& \mathbf{1}^\top \boldsymbol{\eta} = 1, \\
& \boldsymbol{\lambda}, \boldsymbol{\eta} \geq \mathbf{0}, \ \boldsymbol{\pi}_j \geq 0, j \in \mathcal{J} \Big\},
\end{aligned} \tag{13}$$

*where $\mathcal{R}_{\text{NI}} = -\mathbb{R}_+$, $\mathcal{R}_{\text{V}} = \{0\}$, $\mathcal{R}_{\text{ND}} = \mathbb{R}_+$ and $\mathcal{R}_{\text{C}} = \mathbb{R}$.*

Let us clarify the meaning of statement (13). Assume that DMU $(\mathbf{x}, \mathbf{y})$ satisfies conditions in (13) with some vectors $\boldsymbol{\lambda}$, $\boldsymbol{\eta}$ and $\boldsymbol{\pi}_j$, $j \in \mathcal{J}$. It is straightforward to verify that, for each $j'$ such that $\eta_{j'} > 0$, the DMU $(\mathbf{x}, \mathbf{y})$ satisfies all conditions of the statement (9) of the hull $\mathcal{H}^{j}_{\Delta\text{RS-TO}}$ for $j = j'$, if we define $\delta = \lambda_{j'}/\eta_{j'}$ and $\boldsymbol{\pi} = \boldsymbol{\pi}_{j'}/\eta_{j'}$. Therefore, DMU $(\mathbf{x}, \mathbf{y}) \in \mathcal{H}^{j'}_{\Delta\text{RS-TO}}$, for each $j'$ such that $\eta_{j'} > 0$. (It is of course possible that $(\mathbf{x}, \mathbf{y}) \in \mathcal{H}^{j}_{\Delta\text{RS-TO}}$ and $\eta_j = 0$.)

Conversely, let DMU $(\mathbf{x}, \mathbf{y})$ be in the hull $\mathcal{H}^{j'}_{\Delta\text{RS-TO}}$, for some $j' \in \mathcal{J}$, and, therefore, satisfy conditions in (9) stated for $j = j'$, with some scalar $\delta$ and vector $\boldsymbol{\pi}$. Then $(\mathbf{x}, \mathbf{y})$ satisfies all conditions in (13) if we take $\eta_{j'} = \delta$, $\boldsymbol{\pi}_{j'} = \boldsymbol{\pi}$ and $\eta_j = 0$, $\boldsymbol{\pi}_j = \mathbf{0}$, for all $j \in \mathcal{J} \setminus \{j'\}$.

**Remark 4.** In the case of FDH, we have $\mathcal{R}_\Delta = \{0\}$ and, therefore, $\boldsymbol{\eta} = \boldsymbol{\lambda}$. This simplifies the statement (13) of technology $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$ which can be viewed as a special case of the metatechnology developed by

Afsharian and Podinovski (2018):

$$\mathcal{T}^{\rm NC}_{\rm VRS\text{-}TO} = \Big\{ (\mathbf{x},\mathbf{y}) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+ \mid \exists \lambda \in \mathbb{R}^J, \boldsymbol{\pi}_j \in \mathbb{R}^K, j \in \mathcal{J}: \lambda_j \mathbf{x}_j + \sum_{k\in\mathcal{K}} \pi_{jk}\mathbf{p}_k \le \lambda_j \mathbf{x},\ j \in \mathcal{J}, \lambda_j \mathbf{y}_j + \sum_{k\in\mathcal{K}} \pi_{jk}\mathbf{q}_k \ge \lambda_j \mathbf{y},\ j \in \mathcal{J}, \mathbf{1}^\top\lambda = 1, \lambda \ge \mathbf{0}, \boldsymbol{\pi}_j \ge 0, j \in \mathcal{J} \Big\}.$$

## 4. Consistent production trade-offs

In the case of convex multiplier VRS and CRS models, one known possible side effect of weight restrictions is that they may be too restrictive and render the resulting programs infeasible. As shown by Podinovski and Bouzdine-Chameeva (2013, 2015), such outcome is explained by the fact that the technology expanded by the production trade-offs dual to the weight restrictions allows free or unlimited production of a nonzero output vector (we define this formally below.) Podinovski and Bouzdine-Chameeva (2015) call such production trade-offs and corresponding weight restrictions inconsistent with the dataset of observed DMUs. Moreover, they show that inconsistent trade-offs may be undetected by standard calculations of efficiency, in which case the problem with trade-offs remains unknown to the analyst but still results in erroneous efficiency scores. In summary, Podinovski and Bouzdine-Chameeva (2013, 2015) argue that it is good practice to check the consistency of trade-offs in any application, for which they develop analytical and programming tests.

In this section, we consider the notion of consistent trade-offs in nonconvex production technologies $\mathcal{T}^{\rm NC}_{\Delta\rm RS\text{-}TO}$ and develop exact and sufficient tests for checking their consistency.

We start with formal definitions. Given a nonzero output vector $\mathbf{y}_o \in \mathbb{R}^s_+\backslash\{\mathbf{0}\}$, technology $\mathcal{T}$ allows *free* production of $\mathbf{y}_o$ if $(\mathbf{0},\mathbf{y}_o) \in \mathcal{T}$, and *unlimited* production of $\mathbf{y}_o$ if there exists some input vector $\mathbf{x}_o \in \mathbb{R}^m_+$, such that $(\mathbf{x}_o, \alpha\mathbf{y}_o) \in \mathcal{T}$ for all $\alpha \ge 0$.

We further say that the trade-offs (2) are *consistent* in technology $\mathcal{T}^{\rm NC}_{\Delta\rm RS\text{-}TO}$ if such technology does not allow free production and does not allow unlimited production (of any nonzero vector of outputs $\mathbf{y}_o \in \mathbb{R}^s_+\backslash\{\mathbf{0}\}$), and are inconsistent otherwise.

**Example 3.** Consider the modified Example 1 in which, instead of the trade-off $(p_1,q_1) = (-1,-2)$, we have specified the trade-off

$$(\tilde{p}_1,\tilde{q}_1) = (-2,-1). \tag{14}$$

The resulting technology $\mathcal{T}^{\rm NC}_{\rm VRS\text{-}TO}$ is shown in Fig. 2 as the overall shaded area below the line $B'BB''$. Note that the point $B'$ represents free production and, therefore, the trade-off $(\tilde{p}_1,\tilde{q}_1)$ is inconsistent with the dataset.

In this example, after the specification of trade-off (14), DMU $A$ becomes inefficient. Its input radial efficiency becomes equal to $0/2 = 0$ and its projection $A^*$ also represents free production, which indicates that an error was made in the specification of this trade-off.

If the observed dataset excluded DMU $A$ and consisted only of DMUs $B$ and $C$, then the fact that the trade-offs are inconsistent would not be detected by the evaluation of efficiency, as both radial targets of DMUs $B$ and $C$, and their efficiency scores, appear unproblematic. However, such scores would not be acceptable because the projection of DMU $C$ is located on the line $B'B$ which is unrealistic. This observation means that the consistency of trade-offs cannot be confirmed just by seemingly unproblematic efficiency scores and needs to be tested. We consider this below.

Note that, in this example, the single trade-off (14) is clearly the cause of free production. However, generally, individual trade-offs may not be problematic in this sense, but taken together, they can be inconsistent. Podinovski and Bouzdine-Chameeva (2013) consider this in detail.

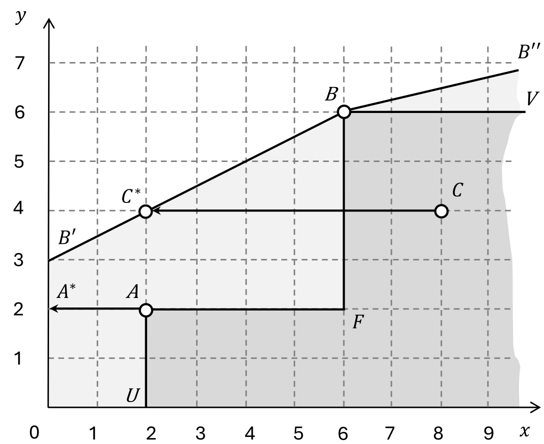


**Fig. 2.** Inconsistent trade-offs generating free production in Example 3.

### 4.1. Mathematical results

In this section, we obtain several mathematical results that clarify the notion of consistent trade-offs and are useful in the development of tests of their consistency.

According to Theorem 3.2, technology $\mathcal{T}^{\rm NC}_{\Delta\rm RS\text{-}TO}$ is the union of hulls $\mathcal{H}^j_{\Delta\rm RS\text{-}TO}$, $j \in \mathcal{J}$. Each of these hulls can be viewed as an elementary convex subtechnology generated by a single observed DMU $(\mathbf{x}_j,\mathbf{y}_j)$ and trade-offs (2), under the assumption of $\Delta$RS. It is clear that the trade-offs (2) are consistent in technology $\mathcal{T}^{\rm NC}_{\Delta\rm RS\text{-}TO}$ if and only if they are consistent in each hull $\mathcal{H}^j_{\Delta\rm RS\text{-}TO}$, $j \in \mathcal{J}$.

This observation allows us to extend the results proved by Podinovski and Bouzdine-Chameeva (2013) in the case of convex technologies, to their nonconvex analogues.

**Theorem 4.1.** *Technology $\mathcal{T}^{\rm NC}_{\rm CRS\text{-}TO}$ allows free production of vector $\mathbf{y}_o$ if and only if it allows its unlimited production.*

It is worth noting that Theorem 4.1 is not true in the cases of VRS, NIRS and NDRS. For example, as shown by Podinovski and Bouzdine-Chameeva (2013), a convex VRS technology (and, similarly, its nonconvex variant $\mathcal{T}^{\rm NC}_{\rm VRS\text{-}TO}$) may allow free production but disallow unlimited production, and vice versa.

**Theorem 4.2.** *Technology $\mathcal{T}^{\rm NC}_{\rm VRS\text{-}TO}$ allows free or unlimited production of vector $\mathbf{y}_o$ if and only if technology $\mathcal{T}^{\rm NC}_{\rm CRS\text{-}TO}$ allows its free and, by Theorem 4.1, its unlimited production. In other words, the trade-offs (2) are consistent or inconsistent in technology $\mathcal{T}^{\rm NC}_{\rm VRS\text{-}TO}$ if and only if they are consistent or, respectively, inconsistent in technology $\mathcal{T}^{\rm NC}_{\rm CRS\text{-}TO}$.*

Another simple but useful observation is that, if technology $\mathcal{T}$ is a subset of technology $\mathcal{T}'$, i.e., $\mathcal{T} \subseteq \mathcal{T}'$, and technology $\mathcal{T}$ allows free or unlimited production, then technology $\mathcal{T}'$ also allows free or, respectively, unlimited production. Note that technology $\mathcal{T}^{\rm NC}_{\rm VRS\text{-}TO}$ is a subset of both technologies $\mathcal{T}^{\rm NC}_{\rm NIRS\text{-}TO}$ and $\mathcal{T}^{\rm NC}_{\rm NDRS\text{-}TO}$, and the two latter technologies are subsets of technology $\mathcal{T}^{\rm NC}_{\rm CRS\text{-}TO}$. This allows us to extend the statement of Theorem 4.2 as follows.

**Theorem 4.3.** *The trade-offs (2) are consistent in any one of the four technologies $\mathcal{T}^{\rm NC}_{\rm VRS\text{-}TO}$, $\mathcal{T}^{\rm NC}_{\rm NIRS\text{-}TO}$, $\mathcal{T}^{\rm NC}_{\rm NDRS\text{-}TO}$ or $\mathcal{T}^{\rm NC}_{\rm CRS\text{-}TO}$ if and only if they are consistent in all of these technologies.*

This last result implies that we do not need to develop tests of consistency of trade-offs for each of the four technologies named in its statement. Instead, it suffices to test their consistency only in technology $\mathcal{T}^{\rm NC}_{\rm CRS\text{-}TO}$. Furthermore, by Theorem 4.1, it suffices to test only if this technology allows unlimited production. This significantly simplifies our task, which is considered below.

### 4.2. Exact test of consistency

As shown in the previous section, the issue of consistency of trade-offs (2) in any of the four nonconvex technologies $\mathcal{T}^{NC}_{\Delta RS\text{-}TO}$ is reduced to their consistency in the technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$. According to Theorem 4.1, it suffices to test if technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$ allows free production of any nonzero output vector $\mathbf{y}_o$.

Clearly, technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$ does not allow free production if and only if the optimal value $y^*$ of the following program is equal to zero, in which $\mathbf{y}$ is a variable vector:

$$y^* = \max \quad \mathbf{1}^\top \mathbf{y} \quad \text{subject to} \quad (\mathbf{0}, \mathbf{y}) \in \mathcal{T}^{NC}_{CRS\text{-}TO}, \quad \mathbf{y} \geq \mathbf{0}. \tag{15}$$

Note that, if program (15) has a nonzero feasible vector $\mathbf{y} \neq \mathbf{0}$ (in which case technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$ allows free production) then, for any $\alpha \geq 0$, the scaled DMU $(\alpha\mathbf{0}, \alpha\mathbf{y}) = (\mathbf{0}, \alpha\mathbf{y}) \in \mathcal{T}^{NC}_{CRS\text{-}TO}$. In this case, in line with Theorem 4.1, technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$ allows unlimited production of vector $\mathbf{y}$ and program (15) has an unbounded objective function.

To solve program (15), we may use any of the statements (8), (10) and (13) of technology $\mathcal{T}^{NC}_{\Delta RS\text{-}TO}$, specified for the case of CRS. However, using the first and last options would result in a nonlinear program, and the statement (10) appears computationally more attractive.

Indeed, because technology $\mathcal{T}^{NC}_{\Delta RS\text{-}TO}$ is the union of hulls $\mathcal{H}^j_{\Delta RS\text{-}TO}$, $j \in \mathcal{J}$, instead of solving program (15), we can test if any of the hulls $\mathcal{H}^j_{\Delta RS\text{-}TO}$, $j \in \mathcal{J}$, allows free production. We state this as the following simple theorem.

**Theorem 4.4.** *The trade-offs (2) are consistent in technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$ if and only if, for each $j \in \mathcal{J}$, the optimal value $y^*_{(j)}$ of the following linear program is equal to zero:*

$$\begin{aligned} y^*_{(j)} = \max \quad & \mathbf{1}^\top \mathbf{y} \\ \text{subject to} \quad & \delta \mathbf{x}_j + \sum_{k \in \mathcal{K}} \pi_k \mathbf{p}_k \leq \mathbf{0}, \\ & \delta \mathbf{y}_j + \sum_{k \in \mathcal{K}} \pi_k \mathbf{q}_k \geq \mathbf{y}, \\ & \delta \geq 0, \ \boldsymbol{\pi} \geq \mathbf{0}, \ \mathbf{y} \geq \mathbf{0}. \end{aligned} \tag{16}$$

According to Theorem 4.4, to test the consistency of trade-offs (2), it suffices to solve linear programs (16) for all $j \in \mathcal{J}$ and verify that all their optimal values are finite and equal to zero. If program (16) has an unbounded objective function for at least one $j \in \mathcal{J}$, then technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$ allows free and unlimited production and the trade-offs (2) are inconsistent.

### 4.3. Sufficient test of consistency

The nonconvex CRS technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$ is a subset of the convex CRS technology $\mathcal{T}^{C}_{CRS\text{-}TO}$. If the former technology allows free or unlimited production, then the latter also allows it. This implies that, if the trade-offs (2) are consistent in the convex technology $\mathcal{T}^{C}_{CRS\text{-}TO}$, then they are also consistent in the nonconvex technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$.

Similar to the approach of Podinovski and Bouzdine-Chameeva (2013), testing consistency of trade-offs (2) in technology $\mathcal{T}^{C}_{CRS\text{-}TO}$ can be done by solving a single linear program. Namely, the trade-offs (2) are consistent in technology $\mathcal{T}^{C}_{CRS\text{-}TO}$ if and only if the optimal value $\hat{y}$ of the following linear program is equal to zero:

$$\hat{y} = \max \quad \mathbf{1}^\top \mathbf{y} \quad \text{subject to} \quad (\mathbf{0}, \mathbf{y}) \in \mathcal{T}^{C}_{CRS\text{-}TO}, \quad \mathbf{y} \geq \mathbf{0}.$$

This program is restated in the extended form as follows:

$$\begin{aligned} \hat{y} = \max \quad & \mathbf{1}^\top \mathbf{y} \\ \text{subject to} \quad & \sum_{j \in \mathcal{J}} \lambda_j \mathbf{x}_j + \sum_{k \in \mathcal{K}} \pi_k \mathbf{p}_k \leq \mathbf{0}, \\ & \sum_{j \in \mathcal{J}} \lambda_j \mathbf{y}_j + \sum_{k \in \mathcal{K}} \pi_k \mathbf{q}_k \geq \mathbf{y}, \\ & \mathbf{y} \geq \mathbf{0}, \ \boldsymbol{\lambda} \geq \mathbf{0}, \ \boldsymbol{\pi} \geq \mathbf{0}. \end{aligned} \tag{17}$$

The next example shows that the consistency of trade-offs in the convex technology $\mathcal{T}^{C}_{CRS\text{-}TO}$ is only a sufficient, but not necessary, condition of their consistency in technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$. Therefore, if program (17) has an unbounded objective function, this does not necessarily mean that the objective function of program (15) is also unbounded and that, equivalently, trade-offs (2) are inconsistent in technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$.

**Example 4.** Consider a modified setting used in Example 2, with the same two observed DMUs $(\mathbf{x}_1, \mathbf{y}_1) = (2, 0, 10, 1)^\top$ and $(\mathbf{x}_2, \mathbf{y}_2) = (0, 2, 10, 1)^\top$, where the first two components are inputs and the last two components are outputs. Let the single trade-off be $(\mathbf{p}, \mathbf{q}) = (-1, -1, -5, -1)^\top$.

Computations show that, in the described setting, the optimal values of both programs (16) stated for $j = 1, 2$ are equal to zero. However, the objective function of program (17) is unbounded. Therefore, the trade-off $(\mathbf{p}, \mathbf{q})$ is consistent in all nonconvex technologies $\mathcal{T}^{NC}_{\Delta RS\text{-}TO}$, including its CRS and VRS variants, but is inconsistent in the convex technology $\mathcal{T}^{C}_{CRS\text{-}TO}$ and, as proved by Podinovski and Bouzdine-Chameeva (2013), in technology $\mathcal{T}^{C}_{VRS\text{-}TO}$. (Similar to Theorem 4.3, it is straightforward to prove that the trade-off $(\mathbf{p}, \mathbf{q})$ is also inconsistent in the convex NIRS and NDRS variants of the technology.)

As an illustration of inconsistency of trade-off $(\mathbf{p}, \mathbf{q})$, consider the simple average of DMUs $(\mathbf{x}_1, \mathbf{y}_1)$ and $(\mathbf{x}_2, \mathbf{y}_2)$ equal to $(1, 1, 10, 1)^\top$, which is an element of the four (VRS, NIRS, NDRS and CRS) convex technologies $\mathcal{T}^{C}_{\Delta RS\text{-}TO}$. Modifying this DMU by the trade-off $(\mathbf{p}, \mathbf{q})$, we obtain the DMU $(\mathbf{x}', \mathbf{y}') = (0, 0, 5, 0)^\top$. Therefore, all four convex technologies $\mathcal{T}^{C}_{\Delta RS\text{-}TO}$ allow free production of the output vector $\mathbf{y} = (5, 0)^\top$ and the trade-off $(\mathbf{p}, \mathbf{q})$ is inconsistent in each of them.

**Remark 5.** Theoretically, the equality $\hat{y} = 0$ is only sufficient but not necessary for the consistency of trade-offs (2) in the nonconvex technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$. However, in practical applications, we may prefer to use this sufficient condition over the exact consistency test based on solving $J$ linear programs (16), for at least two reasons. First, the former is computationally more straightforward than the latter. Second, as Example 4 shows, it is theoretically possible that the objective function of program (17) is unbounded (informally, $\hat{y} = +\infty$) but the trade-offs (2) are consistent in technology $\mathcal{T}^{NC}_{CRS\text{-}TO}$. However, even in this case, we may still be concerned that such trade-offs are inconsistent in its convex analogue $\mathcal{T}^{C}_{CRS\text{-}TO}$ and decide to re-examine the trade-offs as potentially unreliable.

In the following sections of our paper, we assume that trade-offs (2) are consistent in technology $\mathcal{T}^{NC}_{\Delta RS\text{-}TO}$. This can be tested either by the exact approach discussed in Section 4.2, e.g., by solving $J$ linear programs (16), or by verifying the sufficient condition based on solving the single linear program (17).

## 5. Models based on directional distance function

In this section, we consider the assessment of efficiency of DMUs in technology $\mathcal{T}^{NC}_{\Delta RS\text{-}TO}$ based on the directional distance function defined by Chambers, Chung, and Färe (1998). In the next Section 6, we adapt our results to the input and output radial measures of efficiency.

Consider assessing the efficiency of DMU $(\mathbf{x}_o, \mathbf{y}_o) \in \mathcal{T}^{NC}_{\Delta RS\text{-}TO}$ in the direction defined by the nonzero vector $\mathbf{g} = (\mathbf{g}_x, \mathbf{g}_y) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$. This is achieved by the evaluation of the directional distance function defined as follows:

$$\beta_o = \max \quad \beta \quad \text{subject to} \quad (\mathbf{x}_o - \beta\mathbf{g}_x, \mathbf{y}_o + \beta\mathbf{g}_y) \in \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}, \quad \beta \text{ sign free.} \tag{18}$$

Because, in program (18), the value $\beta = 0$ is always feasible, the optimal value $\beta_o$ of the program (18) is nonnegative. The DMU $(\mathbf{x}_o, \mathbf{y}_o)$ is efficient in the direction of vector $\mathbf{g}$ if $\beta_o = 0$, and is inefficient otherwise. Program (18) identifies the directional projection $(\mathbf{x}_o - \beta_o\mathbf{g}_x, \mathbf{y}_o + \beta_o\mathbf{g}_y)$ of the DMU $(\mathbf{x}_o, \mathbf{y}_o)$ on the boundary of technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$.

Below we consider two computational approaches to the solution of program (18). The first of these, referred to as *the decomposition approach*, is based on the representation of technology (10) as the union of hulls $\mathcal{H}^j_{\Delta\text{RS-TO}}$, $j \in \mathcal{J}$, and requires solving $J$ linear programs, for each DMU $(\mathbf{x}_o, \mathbf{y}_o)$ under the assessment. The second, referred to as *the single-stage approach*, is based on the statement of technology (13) and requires solving a single linear program.

### 5.1. The decomposition approach

By Theorem 3.2, technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ is the union of the hulls $\mathcal{H}^j_{\Delta\text{RS-TO}}$, $j \in \mathcal{J}$. This implies that the optimal value $\beta_o$ of program (18) can be found as the maximum of variable $\beta$ evaluated over all individual hulls $\mathcal{H}^j_{\Delta\text{RS-TO}}$, $j \in \mathcal{J}$, that include the DMU $(\mathbf{x}_o, \mathbf{y}_o)$.

Formally, let DMU $(\mathbf{x}_o, \mathbf{y}_o) \in \mathcal{H}^j_{\Delta\text{RS-TO}}$, $j \in \mathcal{J}$. The efficiency of this DMU in the hull $\mathcal{H}^j_{\Delta\text{RS-TO}}$ is evaluated by the directional distance function

$$\beta_o^{(j)} = \max \quad \beta \quad \text{subject to} \quad (\mathbf{x}_o - \beta\mathbf{g}_x, \mathbf{y}_o + \beta\mathbf{g}_y) \in \mathcal{H}^j_{\Delta\text{RS-TO}}, \quad \beta \text{ sign free.} \tag{19}$$

Taking into account the statement (9) of the hulls $\mathcal{H}^j_{\Delta\text{RS-TO}}$, we restate program (19) as the following linear program:

$$\beta_o^{(j)} = \max \quad \beta \tag{20a}$$

subject to

$$\delta\mathbf{x}_j + \sum_{k\in\mathcal{K}} \pi_k\mathbf{p}_k \leq \mathbf{x}_o - \beta\mathbf{g}_x, \tag{20b}$$

$$\delta\mathbf{y}_j + \sum_{k\in\mathcal{K}} \pi_k\mathbf{q}_k \geq \mathbf{y}_o + \beta\mathbf{g}_y, \tag{20c}$$

$$\mathbf{x}_o - \beta\mathbf{g}_x \geq \mathbf{0}, \tag{20d}$$

$$\mathbf{y}_o + \beta\mathbf{g}_y \geq \mathbf{0}, \tag{20e}$$

$$\delta \in \mathcal{I}_\Delta, \ \boldsymbol{\pi} \geq \mathbf{0}, \ \beta \text{ sign free.} \tag{20f}$$

In the above program, the constraints (20d) and (20e) are a restatement of the condition $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^m_+ \times \mathbb{R}^s_+$ in the statement (9) of the hull $\mathcal{H}^j_{\Delta\text{RS-TO}}$.

As established by Chambers et al. (1998), the directional distance function $\beta_o^{(j)}$ provides a complete characterization of the hull $\mathcal{H}^j_{\Delta\text{RS-TO}}$. Namely, the DMU $(\mathbf{x}_o, \mathbf{y}_o) \in \mathcal{H}^j_{\Delta\text{RS-TO}}$ if and only if program (20) is feasible and $\beta_o^{(j)} \geq 0$. (Because we assume that technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ and therefore any of its hulls $\mathcal{H}^j_{\Delta\text{RS-TO}}$, $j \in \mathcal{J}$, do not allow free and unlimited production, if program (20) is feasible, then its optimal value $\beta_o^{(j)}$ is finite.) This implies that the DMU $(\mathbf{x}_o, \mathbf{y}_o) \notin \mathcal{H}^j_{\Delta\text{RS-TO}}$ if and only if $\beta_o^{(j)} < 0$ or program (20) is infeasible. In the latter case we can formally take $\beta_o^{(j)} = -\infty$.

It is now clear that we have

$$\beta_o = \max\{\beta_o^{(j)} \mid j \in \mathcal{J}\}. \tag{21}$$

In summary, in line with (21), the described approach to the computation of $\beta_o$ consists of solving the linear programs (20) stated for each $j \in \mathcal{J}$, and subsequently obtaining $\beta_o$ as the maximum of the optimal values of those linear programs that have a finite optimal value $\beta_o^{(j)}$. (As agreed, if program (20) is infeasible, we formally take $\beta_o^{(j)} = -\infty$. We can ignore such cases in the computation of $\beta_o$ by formula (21)).

**Remark 6.** The value $\beta_o$ calculated by formula (21) remains unchanged if we solve all programs (20), $j \in \mathcal{J}$, without the constraint (20e). We consider this in detail in Appendix C. In particular, we prove that program (20) has a finite optimal value (nonnegative if the DMU $(\mathbf{x}_o, \mathbf{y}_o) \in \mathcal{H}^j_{\Delta\text{RS-TO}}$ and negative otherwise) if and only if program (20) with the constraint (20e) removed has a final optimal value, in which case the two optimal values are equal. Furthermore, if program (20) is infeasible (i.e., if $(\mathbf{x}_o, \mathbf{y}_o) \notin \mathcal{H}^j_{\Delta\text{RS-TO}}$), then program (20) without its constraint (20e) may have a finite negative optimal value. This does not affect the maximum $\beta_o$ in (21).

### 5.2. The single-stage approach

As an alternative to the decomposition approach that requires solving $J$ linear programs (20), below we consider a single-stage approach that requires solving a single LP, but of a higher dimensionality. (We use this approach in the application in Section 7.)

Based on the statement (13) of technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$, the directional distance function is restated in the following expanded form:

$$\beta_o = \max \quad \beta \tag{22a}$$

subject to

$$\lambda_j\mathbf{x}_j + \sum_{k\in\mathcal{K}} \pi_{jk}\mathbf{p}_k \leq \eta_j\left(\mathbf{x}_o - \beta\mathbf{g}_x\right), \ j \in \mathcal{J}, \tag{22b}$$

$$\lambda_j\mathbf{y}_j + \sum_{k\in\mathcal{K}} \pi_{jk}\mathbf{q}_k \geq \eta_j\left(\mathbf{y}_o + \beta\mathbf{g}_y\right), \ j \in \mathcal{J}, \tag{22c}$$

$$\lambda_j - \eta_j \in \mathcal{R}_\Delta, \ j \in \mathcal{J}, \tag{22d}$$

$$\mathbf{x}_o - \beta\mathbf{g}_x \geq \mathbf{0}, \tag{22e}$$

$$\mathbf{1}^\top\boldsymbol{\eta} = 1, \tag{22f}$$

$$\boldsymbol{\lambda}, \boldsymbol{\eta} \geq \mathbf{0}, \ \boldsymbol{\pi}_j \geq \mathbf{0}, \ j \in \mathcal{J}, \ \beta \text{ sign free.} \tag{22g}$$

Note that $\beta_o \geq 0$ for any DMU $(\mathbf{x}_o, \mathbf{y}_o) \in \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. Therefore, the inequality (20e) would be redundant in model (22) and is not specified. Although program (22) is nonlinear, it can be linearized by the substitution $\beta_j = \eta_j\beta$, $j \in \mathcal{J}$. The resulting linear program is

$$\hat{\beta}_o = \max \quad \sum_{j\in\mathcal{J}} \beta_j \tag{23a}$$

subject to

$$\lambda_j\mathbf{x}_j + \sum_{k\in\mathcal{K}} \pi_{jk}\mathbf{p}_k \leq \eta_j\mathbf{x}_o - \beta_j\mathbf{g}_x, \ j \in \mathcal{J}, \tag{23b}$$

$$\lambda_j\mathbf{y}_j + \sum_{k\in\mathcal{K}} \pi_{jk}\mathbf{q}_k \geq \eta_j\mathbf{y}_o + \beta_j\mathbf{g}_y, \ j \in \mathcal{J}, \tag{23c}$$

$$\lambda_j - \eta_j \in \mathcal{R}_\Delta, \ j \in \mathcal{J}, \tag{23d}$$

$$\eta_j\mathbf{x}_o - \beta_j\mathbf{g}_x \geq \mathbf{0}, \ j \in \mathcal{J}, \tag{23e}$$

$$\mathbf{1}^\top\boldsymbol{\eta} = 1, \tag{23f}$$

$$\boldsymbol{\lambda}, \boldsymbol{\eta} \geq \mathbf{0}, \ \boldsymbol{\pi}_j \geq \mathbf{0}, j \in \mathcal{J}, \ \boldsymbol{\beta} \text{ sign free vector.} \tag{23g}$$

The following theorem is true under the assumption that technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ does not allow free production, which follows from the more general assumption that the trade-offs (2) are consistent with the dataset. (Formally, the next theorem also remains true under the alternative condition that the vector $\mathbf{g}_x \neq \mathbf{0}$.)

**Theorem 5.1.** *Program (22) has an optimal solution if and only if program (23) has an optimal solution, in which case their optimal values are equal:* $\beta_o = \hat{\beta}_o$.

## 6. The input and output radial models

In this section, we consider the assessment of the input and output radial efficiency of the DMU $(\mathbf{x}_o, \mathbf{y}_o)$ in technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. We derive the required models as special cases of the directional distance function approach developed in Section 5.

### 6.1. The input-oriented model

Let $\theta_o$ denote the input radial efficiency of DMU $(\mathbf{x}_o, \mathbf{y}_o)$ in technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. The DMU $(\mathbf{x}_o, \mathbf{y}_o)$ is input radial efficient if and only if $\theta_o = 1$. To evaluate $\theta_o$, we adapt the two approaches developed for the evaluation of directional distance function in Section 5.

The decomposition approach requires solving $J$ linear programs (20) in which we take $\mathbf{g} = (\mathbf{x}_o, \mathbf{0})$ and make the substitution $\theta = 1 - \beta$:

$$\begin{aligned} \theta_o^{(j)} = \min \quad & \theta \\ \text{subject to} \quad & \delta \mathbf{x}_j + \sum_{k\in\mathcal{K}} \pi_k \mathbf{p}_k \le \theta \mathbf{x}_o, \\ & \delta \mathbf{y}_j + \sum_{k\in\mathcal{K}} \pi_k \mathbf{q}_k \ge \mathbf{y}_o, \\ & \delta \in \mathcal{I}_\Delta,\ \boldsymbol{\pi} \ge \mathbf{0},\ \theta \ge 0. \end{aligned} \tag{24}$$

Note that the constraint (20d) of program (20) becomes the non-negativity condition $\theta \ge 0$ in program (24), and the constraint (20e) is always true and is removed.

It is clear that, if $(\mathbf{x}_o, \mathbf{y}_o) \in \mathcal{H}^j_{\Delta\text{RS-TO}}$, program (24) is feasible (with $\theta = 1$) and we have $0 \le \theta_o^{(j)} \le 1$. If $(\mathbf{x}_o, \mathbf{y}_o) \notin \mathcal{H}^j_{\Delta\text{RS-TO}}$, we can formally take $\theta_o^{(j)} = +\infty$. Then the input radial efficiency $\theta_o$ of the DMU $(\mathbf{x}_o, \mathbf{y}_o)$ in technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ is obtained by the variant of formula (21), which is stated as follows:

$$\theta_o = \min\{\theta_o^{(j)} \mid j \in \mathcal{J}\}.$$

The linear programming approach of computing $\theta_o$ is based on solving the linear program (23) in which we take $\mathbf{g} = (\mathbf{x}_o, \mathbf{0})$ and make the variable substitution $\boldsymbol{\theta} = \boldsymbol{\eta} - \boldsymbol{\beta}$. Then $\theta_o$ is equal to the optimal value of the following linear program:

$$\begin{aligned} \theta_o = \min \quad & \sum_{j\in\mathcal{J}} \theta_j \\ \text{subject to} \quad & \lambda_j \mathbf{x}_j + \sum_{k\in\mathcal{K}} \pi_{jk} \mathbf{p}_k \le \theta_j \mathbf{x}_o,\ j \in \mathcal{J}, \\ & \lambda_j \mathbf{y}_j + \sum_{k\in\mathcal{K}} \pi_{jk} \mathbf{q}_k \ge \eta_j \mathbf{y}_o,\ \ j \in \mathcal{J}, \\ & \lambda_j - \eta_j \in \mathcal{R}_\Delta,\ j \in \mathcal{J}, \\ & \mathbf{1}^\top \boldsymbol{\eta} = 1,\ \boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\theta} \ge \mathbf{0},\ \boldsymbol{\pi}_j \ge 0, j \in \mathcal{J}. \end{aligned}$$

### 6.2. The output-oriented model

Let $\psi_o$ denote the output radial efficiency of the DMU $(\mathbf{x}_o, \mathbf{y}_o)$ in technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. The DMU $(\mathbf{x}_o, \mathbf{y}_o)$ is output radial efficient if and only if $\psi_o = 1$. The output radial efficiency of the DMU $(\mathbf{x}_o, \mathbf{y}_o)$ is equal to the inverse of the maximal output improvement factor of this DMU in technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$.

Following the decomposition approach developed in Section 5.1, we solve $J$ linear programs derived from programs (20) by taking $\mathbf{g} = (\mathbf{0}, \mathbf{y}_o)$ and using the variable substitution $\varphi = 1 + \beta$:

$$\begin{aligned} \varphi_o^{(j)} = \max \quad & \varphi \\ \text{subject to} \quad & \delta \mathbf{x}_j + \sum_{k\in\mathcal{K}} \pi_k \mathbf{p}_k \le \mathbf{x}_o, \\ & \delta \mathbf{y}_j + \sum_{k\in\mathcal{K}} \pi_k \mathbf{q}_k \ge \varphi \mathbf{y}_o, \\ & \delta \in \mathcal{I}_\Delta,\ \boldsymbol{\pi} \ge \mathbf{0},\ \varphi \text{ sign free.} \end{aligned} \tag{25}$$

Note that both constraints (20d) and (20e) of program (20) are redundant in program (25) and do not need to be specified.

It is clear that, if $(\mathbf{x}_o, \mathbf{y}_o) \in \mathcal{H}^j_{\Delta\text{RS-TO}}$, program (25) is feasible (with $\varphi = 1$) and we have $\varphi_o^{(j)} \ge 1$. If $(\mathbf{x}_o, \mathbf{y}_o) \notin \mathcal{H}^j_{\Delta\text{RS-TO}}$, we can formally take $\varphi_o^{(j)} = -\infty$. Restating formula (21), we obtain the output radial efficiency of DMU $(\mathbf{x}_o, \mathbf{y}_o)$ as the inverse value

$$\psi_o = 1 / \max\{\varphi_o^{(j)} \mid j \in \mathcal{J}\}.$$

Alternatively, the output radial efficiency $\psi_o$ is equal to the inverse of the optimal value $\varphi_o$ of the following linear program:

$$\begin{aligned} \varphi_o = \max \quad & \sum_{j\in\mathcal{J}} \varphi_j \\ \text{subject to} \quad & \lambda_j \mathbf{x}_j + \sum_{k\in\mathcal{K}} \pi_{jk} \mathbf{p}_k \le \eta_j \mathbf{x}_o,\ j \in \mathcal{J}, \\ & \lambda_j \mathbf{y}_j + \sum_{k\in\mathcal{K}} \pi_{jk} \mathbf{q}_k \ge \varphi_j \mathbf{y}_o,\ \ j \in \mathcal{J}, \\ & \lambda_j - \eta_j \in \mathcal{R}_\Delta,\ j \in \mathcal{J}, \\ & \mathbf{1}^\top \boldsymbol{\eta} = 1, \\ & \boldsymbol{\lambda}, \boldsymbol{\eta} \ge \mathbf{0},\ \boldsymbol{\pi}_j \ge \mathbf{0}, j \in \mathcal{J},\ \boldsymbol{\varphi} \text{ sign free,} \end{aligned} \tag{26}$$

which is derived from the linear program (23) by specifying the direction vector $\mathbf{g} = (\mathbf{0}, \mathbf{y}_o)$ and employing the variable substitution $\boldsymbol{\varphi} = \boldsymbol{\eta} + \boldsymbol{\beta}$.

## 7. Illustrative application

We demonstrate the models and methodologies developed in this paper by an illustrative application to a sample of 126 UK universities.[1]

The data on costs and number of students was obtained from the publicly available website of the Higher Education Statistical Agency (HESA), collected in 2020/21. The data on publications was obtained from the Scopus database for 2020. (Papaioannou and Podinovski (2023) used a similar dataset but, in their study, the costs were disaggregated by department.)

### 7.1. Inputs and outputs

We use a single input (total costs) and seven outputs. The latter include undergraduate (UG) and postgraduate (PG) students, which are accounted separately for the science, non-science (humanities and arts) and medical departments. Such treatment is in line with some literature on higher education (see, e.g., Thanassoulis, Kortelainen, Johnes, & Johnes, 2011) and takes into account that the costs of teaching students at the three groups of departments are generally different. The last output is the number of published papers as a measure of research activity of the universities.

A summary of the data is shown in Table 1.

### 7.2. Specification of production trade-offs

We specify seven production trade-offs which reflect the relative costs of teaching different types of students.

The first three trade-offs state that the teaching of one UG student does not incur more costs to the university than the teaching of one PG student (within the same department), because PG students are generally taught in smaller classes or require more individual supervision than UG students. This means that it should be technologically possible (if a need arises) to reduce the number of PG students at any

[1] This should not be viewed as a fully developed application but rather as a typical context which allows us to discuss the meaning of production trade-offs, their effect on efficiency discrimination and the evaluation of scale properties in the resulting nonconvex models.

**Table 1**
Descriptive statistics for the data in the application.

| Inputs and outputs | Mean | Median | Minimum | Maximum | Standard deviation |
|---|---|---|---|---|---|
| Input 1: Total costs (£1 million) | 279.94 | 204.19 | 17.44 | 2,145.5 | 311.44 |
| Output 1: UG medical students | 390.4 | 0 | 0 | 2,875 | 712.38 |
| Output 2: UG science students | 5,828.49 | 5,257.5 | 200 | 51,040 | 5,011.48 |
| Output 3: UG non-science students | 7,955.44 | 6,687.5 | 50 | 69,400 | 6,649.34 |
| Output 4: PG medical students | 147.58 | 0 | 0 | 2,575 | 363.03 |
| Output 5: PG science students | 1,720.24 | 1,457.5 | 30 | 7,090 | 1,317.92 |
| Output 6: PG non-science students | 2,866.51 | 2,517.5 | 25 | 11,720 | 1,972.42 |
| Output 7: Published papers | 2,063.4 | 866.5 | 9 | 18,495 | 3,259.93 |

department by one and, simultaneously, increase the number of UG students by one, assuming that the total costs and all other types of students remain unchanged.

In line with (2), we state the first three trade-offs as follows:

$$\mathbf{p}_1 = (0),\ \mathbf{q}_1 = (1,0,0,-1,0,0,0)^{\top},$$
$$\mathbf{p}_2 = (0),\ \mathbf{q}_2 = (0,1,0,0,-1,0,0)^{\top}, \qquad (27)$$
$$\mathbf{p}_3 = (0),\ \mathbf{q}_3 = (0,0,1,0,0,-1,0)^{\top}.$$

For example, the first trade-off in (27) means that any medical department would not require more costs if it were to replace one PG student by one UG student, while keeping the other outputs unchanged. The remaining two trade-offs make a similar statement for science and non-science departments.

The next two trade-offs state that the teaching of one UG student at a non-science department is not more costly to the university than the teaching of one UG student at either medical or science department. Therefore, no university should require extra costs (and no change to the other outputs is needed) if it were to increase the non-science students by one and simultaneously reduce the number of medical or science students by one. We state this as the two trade-offs:

$$\mathbf{p}_4 = (0),\ \mathbf{q}_4 = (-1,0,1,0,0,0,0)^{\top},$$
$$\mathbf{p}_5 = (0),\ \mathbf{q}_5 = (0,-1,1,0,0,0,0)^{\top}. \qquad (28)$$

The final two trade-offs restate trade-offs (28) for PG students and have a similar meaning:

$$\mathbf{p}_6 = (0),\ \mathbf{q}_6 = (0,0,0,-1,0,1,0)^{\top},$$
$$\mathbf{p}_7 = (0),\ \mathbf{q}_7 = (0,0,0,0,-1,1,0)^{\top}. \qquad (29)$$

We test the consistency of the stated trade-offs $(\mathbf{p}_k, \mathbf{q}_k)$, $k = 1, \ldots, 7$, by solving an appropriately specified CRS model (17). Its optimal value is equal to zero, which confirms that the specified seven trade-offs are consistent with the dataset.

**Remark 7.** In line with our discussion in Section 1.1, it is worth highlighting that the above production trade-offs are not the same as the exact marginal rates of transformation between different types of students (which are usually unavailable) but are assumed conservative bounds on such rates that should be acceptable to all universities. For example, suppose that, depending on the university, instead of one PG student, non-science departments can use the same resources to teach between 1 and 5 UG students. Then no department can argue that the trade-off $(\mathbf{p}_1, \mathbf{q}_1)$ is technologically unrealistic. Because we do not know the exact range of marginal rates, we make a simplifying assumption that teaching a PG student generally requires at least the same resources as one UG student, and often more. If, in a particular application, this is felt to be an unjustified assumption, we can use a more relaxed trade-off, for example, stating that, instead of 2 PG students, any department should be able to teach one UG student.

### 7.3. Computations and analysis

We evaluate the output radial efficiency of the universities in five different models, all of which are used in the analysis below.

**Table 2**
Output radial efficiency in different models.

| Model | Number of efficient DMUs | Average efficiency | Minimum efficiency | Standard deviation |
|---|---|---|---|---|
| $\mathbb{M}^{NC}_{VRS}$ | 115 | 0.9883 | 0.7807 | 0.0429 |
| $\mathbb{M}^{NC}_{VRS-TO}$ | 95 | 0.9575 | 0.6155 | 0.0882 |
| $\mathbb{M}^{NC}_{CRS-TO}$ | 35 | 0.8371 | 0.4087 | 0.1573 |
| $\mathbb{M}^{NC}_{NIRS-TO}$ | 55 | 0.8887 | 0.4157 | 0.1412 |
| $\mathbb{M}^{NC}_{NDRS-TO}$ | 46 | 0.863 | 0.4087 | 0.1484 |

Model $\mathbb{M}^{NC}_{VRS}$ is the standard FDH model which does not incorporate any of the trade-offs. Model $\mathbb{M}^{NC}_{VRS\text{-}TO}$ is the FDH model in which we incorporate all seven tradeoffs (27)–(29). Models $\mathbb{M}^{NC}_{CRS\text{-}TO}$, $\mathbb{M}^{NC}_{NIRS\text{-}TO}$ and $\mathbb{M}^{NC}_{NDRS\text{-}TO}$ are the CRS, NIRS and NDRS extensions of the model $\mathbb{M}^{NC}_{VRS\text{-}TO}$.

For computations using the models with trade-offs, we use their statement (26). Specifically, for each $j \in \mathcal{J}$, the constraint $\lambda_j - \eta_j \in \mathcal{R}_\Delta$ is changed to $\lambda_j - \eta_j = 0$ in the case of VRS, is removed in the case of CRS and is changed to $\lambda_j - \eta_j \leq 0$ and $\lambda_j - \eta_j \geq 0$ in the cases on NIRS and NDRS, respectively.

Table 2 shows the results of computations using the five models. The first two rows show that the specification of trade-offs noticeably increases the discrimination of the FDH model. Such improvement is achieved by the incorporation of additional information about the relative costs of different types of students stated as production trade-offs.

### 7.4. Scale characteristics

The notions of scale efficiency, MPSS and returns to scale were introduced to DEA by Banker (1984), Banker et al. (1984), and Banker and Thrall (1992). The extension of the notions of scale efficiency and MPSS to nonconvex technologies is unproblematic, but the notion of returns to scale requires a more careful treatment. We consider this below.

The scale efficiency of each university is obtained by dividing its efficiency in model $\mathbb{M}^{NC}_{CRS\text{-}TO}$ by its efficiency in model $\mathbb{M}^{NC}_{VRS\text{-}TO}$. (If a university is not output radial efficient in model $\mathbb{M}^{NC}_{VRS\text{-}TO}$, then its scale efficiency coincides with the scale efficiency of its output projection on the boundary of the FDH technology with trade-offs.) Calculations show that the average scale efficiency across all 126 universities is equal to 0.8719.

Table 2 shows that 35 universities are output radial efficient in the model $\mathbb{M}^{NC}_{CRS\text{-}TO}$. This means that, out of 95 universities that are output radial efficient in the model $\mathbb{M}^{NC}_{VRS\text{-}TO}$, only 35 are at MPSS and their scale efficiency is equal to 1. The remaining $95 - 35 = 60$ universities are not at MPSS, and their scale efficiency is less than 1.

We can now use the approach originally developed by Färe et al. (1983, 1985) and further refined by Podinovski (2004a, 2004b), in order to investigate whether a particular university which is efficient in model $\mathbb{M}^{NC}_{VRS\text{-}TO}$ but whose scale efficiency is less than 1, is smaller or larger than its MPSS. In line with the definition of global returns to scale of Podinovski (2004a), in the former case we say that the university exhibits global increasing returns to scale (G-IRS) and would need to increase its scale of operations to achieve its MPSS. In the latter case, the university exhibits global decreasing returns to scale (G-DRS) and would need to reduce its scale of operations on the way to its MPSS. (As proved by Podinovski (2017), in any convex technology, the notion of global returns to scale is identical to the notion of local returns to scale defined by scale elasticities. In a nonconvex technology, the global and local characterizations are generally different.)

In line with Podinovski (2004a), a university exhibits G-IRS if its output radial efficiency in model $\mathbb{M}^{\text{NC}}_{\text{NIRS-TO}}$ is strictly less than its efficiency in model $\mathbb{M}^{\text{NC}}_{\text{NDRS-TO}}$. The opposite sign between the two efficiency scores identifies universities that exhibit G-DRS. Computations show that, in our sample, 25 universities exhibit G-IRS and 35 exhibit G-DRS.

In summary, out of 95 universities that are output radial efficient in the model $\mathbb{M}^{\text{NC}}_{\text{VRS-TO}}$, 35 are at MPSS, 25 exhibit G-IRS and are smaller than their optimal size, and 35 exhibit G-DRS and are larger than their optimal size.

## 8. Conclusion

In this paper, we considered the incorporation of production trade-offs in the FDH model of technology and its CRS, NIRS and NDRS variants. Such trade-offs are well-established in convex DEA models and represent simultaneous changes to the inputs and outputs that are regarded technologically possible at all DMUs in the technology.

The FDH model is often regarded as the “pure” model of technology in that any DMU whose efficiency is assessed is benchmarked only against the observed DMUs, and not against their convex combinations or any other hypothetical DMUs, as is the case in the standard VRS models. In this sense, the expansion of FDH technology by production trade-offs violates this “purity” and incorporates additional hypothetical DMUs obtained as modifications of observed DMUs by the specified trade-offs used in varying proportions. However, it is worth highlighting that the assumption that certain trade-offs are technologically feasible is unrelated to the assumption of convexity.

Furthermore, we would argue that production trade-offs represent our knowledge of the technology and there is no reason not to specify them as long as we are certain in their validity. As an example we can point at the illustrative application to universities discussed in Section 7. In this application, the trade-offs stated that the teaching of students on non-science degrees is generally not more expensive than teaching them on science and medical degrees, and that similarly teaching undergraduate students is not more costly than teaching postgraduate students. In the authors’ view, such judgements are valid and can be confidently incorporated in the FDH model, the mathematical tools for which are developed in this paper.

We explore the FDH technology with production trade-offs in a single unifying development with its CRS, NIRS and NDRS variants. These variants are often used as reference technologies in the assessment of scale efficiency, MPSS and returns to scale in the underlying FDH technology with trade-offs. We demonstrate the described use of all such technologies in the application in Section 7.

In order to provide a rigorous development of the FDH technology with production trade-offs and its CRS, NIRS and NDRS variants, we utilize the axiomatic approach in which the technology is defined by the minimum extrapolation principle applied to the stated axioms. The immediate operational statements of these technologies are generally nonlinear and utilize binary variables. We address this issue by developing equivalent linearized statements of these technologies. We show that different statements are useful for different purposes, such as theoretical exploration and computations.

We further explore the notion of consistent trade-offs which has previously been developed for convex VRS and CRS models. The trade-offs are inconsistent if the model of technology incorporating them allows free or unlimited production of a nonzero vector of outputs. This indicates that we made an error in the assessment of trade-offs and need to revisit their statements. In this paper, we develop exact and sufficient programming approaches to the testing of consistency of trade-offs.

We finally consider the models for assessment of efficiency of DMUs in the four nonconvex technologies with trade-offs, using the directional distance function approach and also the conventional input and output radial efficiency measures. It is known that the standard FDH models can be solved by computationally efficient enumeration algorithms—see, e.g., Tulkens (1993), Cherchye, Kuosmanen, and Post (2001) and Briec and Kerstens (2006). However, such approaches do not apply to the FDH model extended by production trade-offs and to its CRS, NIRS and NDRS variants. This means that the nonconvex models with trade-offs require a programming solution approach. However, it turns out that, even if we use a linearized statement of the FDH technology with trade-offs or its CRS, NIRS or NDRS analogues, the resulting programs become nonlinear. We further show that all these programs can be linearized.

In Section 7, we provide an application of the methodology and models developed in this paper to the assessment of efficiency of UK universities. In particular, we discuss the statement of several production trade-offs in the given context and further show that their specification significantly improves the discriminating power of the FDH model and its CRS, NIRS and NDRS analogues. We also discuss the evaluation of scale efficiency and returns to scale in the given setting.

### CRediT authorship contribution statement

**Mahmood Mehdiloo:** Writing – original draft, Writing – review & editing. **Grammatoula Papaioannou:** Writing – original draft, Writing – review & editing. **Victor V. Podinovski:** Writing – original draft, Writing – review & editing.

## Appendix A. Axiomatic variations

In this section, we provide additional details for the discussion in Remark 1.

Podinovski (2004d) makes an assumption of feasibility of *individual* production trade-offs (2) by stating the following axiom. (We adjust the original notation in line with the notation used in our paper.)

**Axiom FITO** (*Feasibility of Individual Trade-Offs*) If DMU $(\mathbf{x},\mathbf{y}) \in \mathcal{T}$ then, for any $k \in \mathcal{K}$ and any $\pi_k \geq 0$, the DMU $(\mathbf{x}',\mathbf{y}') = (\mathbf{x},\mathbf{y}) + \pi_k(\mathbf{p}_k,\mathbf{q}_k) \in \mathcal{T}$, provided $\mathbf{x} + \pi_k\mathbf{p}_k \geq \mathbf{0}$ and $\mathbf{y} + \pi_k\mathbf{q}_k \geq \mathbf{0}$.

It may appear that the original Axiom FITO is equivalent to Axiom FTO stated in Section 2.3 which allows simultaneous modifications of the DMU $(\mathbf{x},\mathbf{y})$ by a combination of all production trade-offs. This is because, apparently, the trade-offs in such combination can be used consecutively, one after another, resulting in the same DMU $(\mathbf{x}',\mathbf{y}')$ as defined by (3). This appearance is, however, generally incorrect.

Indeed, consider a DMU $(\mathbf{x},\mathbf{y})$ with three outputs stated by the vector $\mathbf{y} = (0,0,5)^\top$. (Note that output vectors with zero components are always present in a technology that is strongly disposable in the outputs.) Further consider the two trade-offs for the outputs stated by the vectors $\mathbf{q}_1 = (-1,2,-1)^\top$ and $\mathbf{q}_2 = (2,-1,-1)^\top$.

Neither of the trade-offs $\mathbf{q}_1$ nor $\mathbf{q}_2$ can individually be used to modify the DMU $(\mathbf{x},\mathbf{y})$ because doing so would result in a negative output 1 or 2. Therefore, in the described scenario, the original Axiom FITO does not generate any new DMUs. However, using these two trade-offs in combination with, e.g., $\pi_1 = \pi_2 = 1$, equal to $\mathbf{q}_1 + \mathbf{q}_2 = (1,1,-2)^\top$, avoids creating negative outputs and can, by Axiom FTO, be used to modify the DMU $(\mathbf{x},\mathbf{y})$. More precisely, the resulting DMU $(\mathbf{x}',\mathbf{y}')$ has the output vector

$$\mathbf{y}' = \mathbf{y} + \pi_1\mathbf{q}_1 + \pi_2\mathbf{q}_2 = (1,1,3)^\top.$$

This example shows that Axiom FTO is a stronger assumption than the original Axiom FITO. For the same set of production trade-offs (2), the former may generate DMUs $(\mathbf{x}',\mathbf{y}')$ that the latter cannot.

In the proof of Theorems 1 and 2 of Podinovski (2004d), which establish an explicit statement of the CRS and VRS technologies expanded by production trade-offs, this "weakness" of Axiom FITO is compensated by the additional axiom that the technology is a closed set. (This standard assumption should be true anyway and either required by an additional axiom or be a property following from the other axioms.) A further required assumption is that the technology includes at least one DMU $(\mathbf{x}^*, \mathbf{y}^*)$ whose all inputs and outputs are strictly positive. (A sufficient condition of this for the CRS or VRS technology is that every input and output is strictly positive in at least one observed DMU. In this case, the required DMU $(\mathbf{x}^*, \mathbf{y}^*)$ can be defined as the simple average of all observed DMUs.)

In axiomatic definitions, the goal is to define a technology by the weakest (and therefore most reliable) set of assumptions. In this sense, for defining technology $\mathcal{T}^{\text{C}}_{\Delta\text{RS}}$, the use of the weaker original Axiom FITO over the stronger Axiom FTO is preferable. However, the difference between the two axioms is almost imperceptible and using either axiom leads to the same statement of technology. In this paper we use the stronger Axiom FTO, which helps us to avoid mathematical complications arising from the use of Axiom FITO.

## Appendix B. Proofs

**Proof of Theorem 3.1.** Denote $\mathcal{T}'$ the technology (set) defined by the conditions in the braces on the right-hand side of (8). We need to prove that $\mathcal{T}' = \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. Clearly, technology $\mathcal{T}'$ satisfies Axioms IO and SD. To prove that $\mathcal{T}'$ satisfies Axiom $\Delta$RS, consider any DMU $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}'$. By definition, $(\mathbf{x}, \mathbf{y})$ satisfies conditions in (8) with some vectors $\lambda'$, $\boldsymbol{\pi}'$ and scalar $\delta' \in \mathcal{I}_\Delta$. Consider any $\delta = \delta^* \in \mathcal{I}_\Delta$ in the statement of Axiom $\Delta$RS. Then the DMU $(\delta^*\mathbf{x}, \delta^*\mathbf{y})$ satisfies the conditions in (8) with $\lambda = \lambda'$, $\boldsymbol{\pi} = \delta^*\boldsymbol{\pi}'$ and scalar $\delta = \delta^*\delta' \in \mathcal{I}_\Delta$. Therefore, $(\delta^*\mathbf{x}, \delta^*\mathbf{y}) \in \mathcal{T}'$. This implies that $\delta^*\mathcal{T}' \subseteq \mathcal{T}'$ for all $\delta^* \in \mathcal{I}_\Delta$. Therefore, technology $\mathcal{T}'$ satisfies Axiom $\Delta$RS.

To prove that $\mathcal{T}'$ satisfies Axiom FTO, consider any DMU $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}'$, which satisfies conditions in (8) with some vectors $\lambda'$, $\boldsymbol{\pi}'$ and scalar $\delta' \in \mathcal{I}_\Delta$. Consider any vector $\boldsymbol{\pi} = \boldsymbol{\pi}^*$ in the statement of Axiom FTO. Then the DMU $(\mathbf{x}', \mathbf{y}')$ defined by (3) satisfies conditions in (8) with $\lambda = \lambda'$, $\boldsymbol{\pi} = \boldsymbol{\pi}' + \boldsymbol{\pi}^*$ and scalar $\delta = \delta'$. Therefore, technology $\mathcal{T}'$ satisfies Axiom FTO. This implies that $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}} \subseteq \mathcal{T}'$.

Conversely, consider any DMU $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}'$ which satisfies all conditions of (8) with some vectors $\lambda'$, $\boldsymbol{\pi}'$ and scalar $\delta'$. To be specific, let $\lambda'_1 = 1$ (and all the other components of vector $\lambda'$ be equal to zero). Then we have

$$\delta'\mathbf{x}_1 + \sum_{k\in\mathcal{K}} \pi'_k \mathbf{p}_k \le \mathbf{x}, \tag{30a}$$

$$\delta'\mathbf{y}_1 + \sum_{k\in\mathcal{K}} \pi'_k \mathbf{q}_k \ge \mathbf{y}. \tag{30b}$$

Because technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ satisfies Axioms IO and $\Delta$RS, the DMU $(\delta'\mathbf{x}_1, \delta'\mathbf{y}_1) \in \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. It may appear that we can now use Axiom FTO to conclude that $(\delta'\mathbf{x}_1 + \sum_{k\in\mathcal{K}} \pi'_k\mathbf{p}_k, \delta'\mathbf{y}_1 + \sum_{k\in\mathcal{K}} \pi'_k\mathbf{q}_k) \in \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. However, the input vector $\delta'\mathbf{x}_1 + \sum_{k\in\mathcal{K}} \pi'_k\mathbf{p}_k$ may have negative components in which case Axiom FTO does not apply. (Note that, by (30b), we always have $\delta'\mathbf{y}_1 + \sum_{k\in\mathcal{K}} \pi'_k\mathbf{q}_k \ge \mathbf{0}$.) To overcome this possibility, we proceed in a different way.

Define the slack vector $\mathbf{s} = \mathbf{x} - (\delta'\mathbf{x}_1 + \sum_{k\in\mathcal{K}} \pi'_k\mathbf{p}_k) \ge \mathbf{0}$. By (30a), we have $\mathbf{s} \ge \mathbf{0}$. By Axiom SD, the DMU $(\delta'\mathbf{x}_1 + \mathbf{s}, \delta'\mathbf{y}_1) \in \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. We now define the DMU $(\mathbf{x}', \mathbf{y}') = (\delta'\mathbf{x}_1 + \mathbf{s} + \sum_{k\in\mathcal{K}} \pi'_k\mathbf{p}_k, \delta'\mathbf{y}_1 + \sum_{k\in\mathcal{K}} \pi'_k\mathbf{q}_k)$. Note that, by the definition of vector $\mathbf{s}$ and by the inequality in (30b), the input and output vectors $\mathbf{x}'$ and $\mathbf{y}'$ of this DMU are nonnegative, as required by Axiom FTO. (We have $\mathbf{x}' = \mathbf{x} \ge \mathbf{0}$ and $\mathbf{y}' \ge \mathbf{y} \ge \mathbf{0}$). Because $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ satisfies Axiom FTO, the DMU $(\mathbf{x}', \mathbf{y}') = (\mathbf{x}, \mathbf{y}') \in \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$.

Taking into account (30b), by Axiom SD, the DMU $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. Therefore, $\mathcal{T}' \subseteq \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. Combining this with the opposite embedding proved above, we obtain $\mathcal{T}' = \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. □

**Proof of Theorem 3.2.** Based on the statement (8), for each $j' \in \mathcal{J}$, the hull $\mathcal{H}^{j'}_{\Delta\text{RS-TO}}$ is the subset of technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ corresponding to the vector $\lambda$ in which $\lambda_{j'} = 1$ and $\lambda_j = 0$, $j \in \mathcal{J} \setminus \{j'\}$. Therefore, $\bigcup_{j\in\mathcal{J}} \mathcal{H}^j_{\Delta\text{RS-TO}} \subseteq \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. Conversely, any DMU $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ satisfies all conditions of the statement (8) with some vectors $\lambda$, $\boldsymbol{\pi}$ and scalar $\delta$. Because components of vector $\lambda$ are binary and $\mathbf{1}^\top\lambda = 1$, there exists a $j' \in \mathcal{J}$ such that $\lambda_{j'} = 1$ and $\lambda_j = 0$, for all $j \in \mathcal{J} \setminus \{j'\}$. Then the DMU $(\mathbf{x}, \mathbf{y})$ satisfies all conditions of the statement (9) with the same vector $\boldsymbol{\pi}$ and scalar $\delta$. Therefore, $(\mathbf{x}, \mathbf{y}) \in \mathcal{H}^{j'}_{\Delta\text{RS-TO}}$, which implies that $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}} \subseteq \bigcup_{j\in\mathcal{J}} \mathcal{H}^j_{\Delta\text{RS-TO}}$. □

**Proof of Theorem 3.3.** Consider any $j \in \mathcal{J}$. Then the fact that $\mathcal{H}^j_{\Delta\text{RS-TO}}$ is a polyhedral set follows from Podinovski et al. (2016). □

**Proof of Theorem 3.4.** Let DMU $(x_o, y_o) \in \mathcal{T}^{\text{C}}_{\text{VRS-TO}}$. Taking into account (12), it suffices to prove that $(x_o, y_o) \in \text{conv}\left(\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}\right)$. Consider the program $\max\{\delta \mid (x_o, (y_o - 1) + \delta) \in \mathcal{T}^{\text{C}}_{\text{VRS-TO}}\}$, which can be restated as the following canonical linear program:

$$\begin{aligned}
\delta^* = \max \quad & \delta \\
\text{subject to} \quad & \sum_{j\in\mathcal{J}} \lambda_j x_j + \sum_{k\in\mathcal{K}} \pi_k p_k + s_x = x_o, \\
& \sum_{j\in\mathcal{J}} \lambda_j y_j + \sum_{k\in\mathcal{K}} \pi_k q_k - s_y = y_o - 1 + \delta, \\
& \mathbf{1}^\top\lambda = 1, \\
& \lambda, \boldsymbol{\pi} \ge \mathbf{0}, s_x, s_y, \delta \ge 0.
\end{aligned} \tag{31}$$

Because technology $\mathcal{T}^{\text{C}}_{\text{VRS-TO}}$ does not allow unlimited production of output, the objective function of program (31) is bounded above. Then this program has a finite optimal solution $(\lambda^*, \boldsymbol{\pi}^*, s^*_x, s^*_y, \delta^*)$ that is a basic feasible solution (BFS). Because the constraints of program (31) are stated as three equalities, this BFS has no more than three strictly positive components. One of them is $\delta^* \ge 1$. We now have two possibilities.

(1) Let $\boldsymbol{\pi}^* = \mathbf{0}$. From (31), $(x_o, y_o) = (\sum_{j\in\mathcal{J}} \lambda^*_j x_j + s^*_x, \sum_{j\in\mathcal{J}} \lambda^*_j y_j - s^*_y - (\delta^* - 1))$. Because $\delta^* - 1 \ge 0$, we have $(x_o, y_o) \in \mathcal{T}^{\text{C}}_{\text{VRS}} = \text{conv}\left(\mathcal{T}^{\text{NC}}_{\text{VRS}}\right) \subseteq \text{conv}\left(\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}\right)$.

(2) Let $\boldsymbol{\pi}^* \ne \mathbf{0}$. Then exactly one component $\lambda^*_{j'}$ of vector $\lambda^*$ is strictly positive and hence equal to 1, exactly one component $\pi^*_{k'}$ of vector $\boldsymbol{\pi}^*$ is strictly positive, and both $s^*_x = s^*_y = 0$. Then, from (31), we have $(x_o, y_o) = (x_{j'} + \pi^*_{k'}p_{k'}, y_{j'} + \pi^*_{k'}q_{k'} - (\delta^* - 1))$. Because $\delta^* - 1 \ge 0$, we have $(x_o, y_o) \in \mathcal{H}^{j'}_{\text{VRS-TO}} \subseteq \text{conv}\left(\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}\right)$. □

**Proof of Theorem 3.5.** Let $\mathcal{H}$ and $\mathcal{G}$ denote the sets on the right-hand side of (10) and (13), respectively. By Theorem 3.2, we need to prove that $\mathcal{H} = \mathcal{G}$.

Let $(\mathbf{x}, \mathbf{y}) \in \mathcal{H}$. By (10), there exists a $j' \in \mathcal{J}$ such that $(\mathbf{x}, \mathbf{y}) \in \mathcal{H}^{j'}_{\Delta\text{RS-TO}}$. Then it follows from (9) that there exist a scalar $\delta \in \mathcal{I}_\Delta$ and a vector $\boldsymbol{\pi} \in \mathbb{R}^K_+$ such that

$$\delta\mathbf{x}_{j'} + \sum_{k\in\mathcal{K}} \pi_k\mathbf{p}_k \le \mathbf{x}, \quad \delta\mathbf{y}_{j'} + \sum_{k\in\mathcal{K}} \pi_k\mathbf{q}_k \ge \mathbf{y}.$$

Define $\hat{\lambda}_{j'} = \delta$, $\hat{\boldsymbol{\pi}}_{j'} = \boldsymbol{\pi}$, $\hat{\eta}_{j'} = 1$, $\hat{\lambda}_j = 0$, $\hat{\boldsymbol{\pi}}_j = \mathbf{0}$ and $\hat{\eta}_j = 0$ for all $j \in \mathcal{J}\setminus\{j'\}$. Then, $(\mathbf{x}, \mathbf{y})$ satisfies (13) with the vectors $\hat{\lambda}$, $\hat{\boldsymbol{\pi}}_j$, $j \in \mathcal{J}$, and $\hat{\boldsymbol{\eta}}$. Therefore, $(\mathbf{x}, \mathbf{y}) \in \mathcal{G}$, and $\mathcal{H} \subseteq \mathcal{G}$.

Conversely, let $(\mathbf{x}, \mathbf{y}) \in \mathcal{G}$. Then, $(\mathbf{x}, \mathbf{y})$ satisfies (13) with some vectors $\lambda$, $\boldsymbol{\eta}$ and $\boldsymbol{\pi}_j$, $j \in \mathcal{J}$. Because $\boldsymbol{\eta} \ge \mathbf{0}$ and $\mathbf{1}^\top\boldsymbol{\eta} = 1$, there exists at least one $j'' \in \mathcal{J}$ such that $0 < \eta_{j''} \le 1$. This implies that

$$\frac{\lambda_{j''}}{\eta_{j''}}\mathbf{x}_{j''} + \sum_{k\in\mathcal{K}} \frac{\pi_{j''k}}{\eta_{j''}}\mathbf{p}_k \le \mathbf{x},$$

$$\frac{\lambda_{j''}}{\eta_{j''}}\mathbf{y}_{j''} + \sum_{k\in\mathcal{K}} \frac{\pi_{j''k}}{\eta_{j''}}\mathbf{p}_k \ge \mathbf{y},$$

$$\frac{\lambda_{j''}}{\eta_{j''}} - 1 \in \mathcal{R}_\Delta.$$

Define $\hat{\delta} = \lambda_{j''}/\eta_{j''}$ and $\hat{\boldsymbol{\pi}} = \boldsymbol{\pi}_{j''}/\eta_{j''}$. Then $(\mathbf{x}, \mathbf{y})$ satisfies (9) with $\hat{\delta}$ and $\hat{\boldsymbol{\pi}}$. This implies $(\mathbf{x}, \mathbf{y}) \in \mathcal{H}$, and thus $\mathcal{G} \subseteq \mathcal{H}$. Therefore, $\mathcal{H} = \mathcal{G}$. □

**Proof of Theorem 4.1.** Because the cone technology $\mathcal{T}^{\text{NC}}_{\text{CRS-TO}}$ is closed, the proof follows from statement (b) of Theorem 2 in Podinovski and Bouzdine-Chameeva (2013). □

**Proof of Theorem 4.2.** If technology $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$ allows free or unlimited production of the vector $\mathbf{y}_o$, then there exists a $j \in \mathcal{J}$ such that the hull $\mathcal{H}^j_{\text{VRS-TO}}$ allows free or, respectively, unlimited production of vector $\mathbf{y}_o$. The hull $\mathcal{H}^j_{\text{VRS-TO}}$ can be viewed as the convex VRS technology generated by the single observed DMU $(\mathbf{x}_j, \mathbf{y}_j)$ and expanded by the trade-offs (2). Then, as proved by Podinovski and Bouzdine-Chameeva (2013), the cone $\mathcal{H}^j_{\text{CRS-TO}}$, and therefore technology $\mathcal{T}^{\text{NC}}_{\text{CRS-TO}}$ as the union of all such cones, allow free and unlimited production of vector $\mathbf{y}_o$. The converse is proved in a similar way. It suffices to note that, as proved by Podinovski and Bouzdine-Chameeva (2013), if the CRS cone $\mathcal{H}^j_{\text{CRS-TO}}$ allows free and unlimited production of vector $\mathbf{y}_o$, then the VRS hull $\mathcal{H}^j_{\text{VRS-TO}}$ and, therefore, technology $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$ as the union of all such hulls, allow its free or unlimited production. □

**Proof of Theorem 4.3.** Suppose that any of the technologies $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$, $\mathcal{T}^{\text{NC}}_{\text{NIRS-TO}}$ or $\mathcal{T}^{\text{NC}}_{\text{NDRS-TO}}$ allows free or unlimited production of a nonzero vector $\mathbf{y}_o$. Because each of these technologies is a subset of technology $\mathcal{T}^{\text{NC}}_{\text{CRS-TO}}$, the latter technology also allows free or unlimited production of vector $\mathbf{y}_o$. (By Theorem 4.1, it allows both free and unlimited production of this vector.) Then, by Theorem 4.2, technology $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$ allows either free or unlimited production of vector $\mathbf{y}_o$. Because technology $\mathcal{T}^{\text{NC}}_{\text{VRS-TO}}$ is a subset of both technologies $\mathcal{T}^{\text{NC}}_{\text{NIRS-TO}}$ and $\mathcal{T}^{\text{NC}}_{\text{NDRS-TO}}$, both of them allow either free or unlimited production of vector $\mathbf{y}_o$, respectively.

In summary, if any of the four technologies allows free or unlimited production, then all of them allow free or unlimited production. Therefore, if the trade-offs (2) are consistent or inconsistent in any of these technologies, they are consistent or, respectively, inconsistent in all of them. □

**Proof of Theorem 5.1.** Let $\beta^*$, $\lambda^*$, $\boldsymbol{\eta}^*$ and $\boldsymbol{\pi}^*_j$, $j \in \mathcal{J}$, be a feasible solution to program (22). Define $\hat{\boldsymbol{\beta}} = \boldsymbol{\eta}^*\beta^*$, $\hat{\lambda} = \lambda^*$, $\hat{\boldsymbol{\eta}} = \boldsymbol{\eta}^*$ and $\hat{\boldsymbol{\pi}}_j = \boldsymbol{\pi}^*_j$, $j \in \mathcal{J}$. It is straightforward to verify that this defines a feasible solution to program (23). In particular, for each $j \in \mathcal{J}$, the inequality (23e) is obtained from the inequality (22e) stated for $\beta^*$, by multiplying its both sides by $\eta^*_j \geq 0$ and substituting $\eta^*_j\beta^*$ by $\hat{\beta}_j$. Because $\mathbf{1}^\top\boldsymbol{\eta}^* = 1$, for this feasible solution, the objective function of program (23) is equal to $\mathbf{1}^\top\hat{\boldsymbol{\beta}} = (\mathbf{1}^\top\boldsymbol{\eta}^*)\beta^* = \beta^*$. Therefore, $\beta_o \leq \hat{\beta}_o$.

Conversely, let $\hat{\boldsymbol{\beta}}$, $\hat{\lambda}$, $\hat{\boldsymbol{\eta}}$ and $\hat{\boldsymbol{\pi}}_j$, $j \in \mathcal{J}$, be a feasible solution to program (23). Consider any $j' \in \mathcal{J}$ such that $\hat{\eta}_{j'} > 0$. (By (23f), at least one such $j'$ exists.) Stating constraints (23b) and (23c) for this solution and dividing their both sides by $\hat{\eta}_{j'}$, we obtain

$$\begin{aligned} &\frac{\hat{\lambda}_{j'}}{\hat{\eta}_{j'}}\mathbf{x}_j + \sum_{k\in\mathcal{K}} \frac{\hat{\pi}_{j'k}}{\hat{\eta}_{j'}}\mathbf{p}_k \leq \frac{\hat{\eta}_{j'}}{\hat{\eta}_{j'}}\mathbf{x}_o - \frac{\hat{\beta}_{j'}}{\hat{\eta}_{j'}}\mathbf{g}_x, \\ &\frac{\hat{\lambda}_{j'}}{\hat{\eta}_{j'}}\mathbf{y}_j + \sum_{k\in\mathcal{K}} \frac{\hat{\pi}_{j'k}}{\hat{\eta}_{j'}}\mathbf{q}_k \geq \frac{\hat{\eta}_{j'}}{\hat{\eta}_{j'}}\mathbf{y}_o + \frac{\hat{\beta}_{j'}}{\hat{\eta}_{j'}}\mathbf{g}_y. \end{aligned} \tag{32}$$

Define $\beta^* = \hat{\beta}_{j'}/\hat{\eta}_{j'}$, $\lambda^*_{j'} = \hat{\lambda}_{j'}/\hat{\eta}_{j'}$, $\boldsymbol{\pi}^*_{j'} = \hat{\boldsymbol{\pi}}_{j'}/\hat{\eta}_{j'}$, $\eta^*_{j'} = 1$. Further define $\lambda^*_j = 0$, $\boldsymbol{\pi}^*_j = \mathbf{0}$ and $\eta^*_j = 0$, for all $j \in \mathcal{J}\backslash\{j'\}$. Then $\beta^*$, $\lambda^*$, $\boldsymbol{\eta}^*$ and $\boldsymbol{\pi}^*_j$, $j \in \mathcal{J}$, is a feasible solution to program (22).

Indeed, for $j = j'$, the inequalities (22b) and (22c) follow from (32). For all $j \in \mathcal{J}\backslash\{j'\}$, the inequalities (22b) and (22c) are trivially satisfied because both their sides are equal to zero. The inequality (23d) stated for $\hat{\lambda}_j$ and $\hat{\eta}_j$ implies the inequality (22d) stated for $\lambda^*_j$ and $\eta^*_j$, for all $j \in \mathcal{J}$. To prove (22e), divide both sides of the inequality (23e) stated for $\hat{\eta}_{j'}$ and $\hat{\beta}_{j'}$ by $\hat{\eta}_{j'} > 0$. For this feasible solution of program (22), its objective function is equal to $\beta^*$. Therefore, we have

$$\beta_o \geq \beta^* = \hat{\beta}_{j'}/\hat{\eta}_{j'}.$$

We have shown that, for any $j = j'$ such that $\hat{\eta}_j > 0$, we have $\beta_o \geq \hat{\beta}_j/\hat{\eta}_j$. Multiply both sides of this inequality by $\hat{\eta}_j > 0$ and sum them over $j$ such that $\hat{\eta}_j > 0$. Because $\mathbf{1}^\top\hat{\boldsymbol{\eta}} = 1$, we have

$$\beta_o \geq \sum_{j\in\mathcal{J}:\, \hat{\eta}_j>0} \hat{\beta}_j. \tag{33}$$

To complete the proof, we need to prove that $\hat{\eta}_j = 0$ implies $\hat{\beta}_j = 0$. We can then remove the condition $\hat{\eta}_j > 0$ under the summation in (33). Let $\hat{\eta}_j = 0$. If $\mathbf{g}_x \neq \mathbf{0}$, then the equality $\hat{\beta}_j = 0$ follows from (23e). If $\mathbf{g}_x = \mathbf{0}$, then $\mathbf{g}_y \neq \mathbf{0}$ and the terms on the right-hand side of inequalities (23b) and (23c) are equal to $\mathbf{0}$ and $\hat{\beta}_j\mathbf{g}_y$, respectively. According to the statement (13), we have $(\mathbf{0}, \hat{\beta}_j\mathbf{g}_y) \in \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$. As assumed, technology $\mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$ does not allow free production. This implies that $\hat{\beta}_j = 0$.

We have proved that the equality $\hat{\eta}_j = 0$ implies $\hat{\beta}_j = 0$. Therefore, we can restate the inequality (33) as $\beta_o \geq \sum_{j\in\mathcal{J}} \hat{\beta}_j$. Because the feasible solution $\hat{\boldsymbol{\beta}}$, $\hat{\lambda}$, $\hat{\boldsymbol{\eta}}$ and $\hat{\boldsymbol{\pi}}_j$, $j \in \mathcal{J}$, of program (23) is arbitrary, we have $\beta_o \geq \hat{\beta}_o$. Combining this inequality with the opposite inequality established in the first part of the proof, we have $\beta_o = \hat{\beta}_o$. □

## Appendix C. Redundancy of constraint (20e)

Recall that program (20) has a finite optimal value $\beta_o \geq 0$ if and only if $(\mathbf{x}_o, \mathbf{y}_o) \in \mathcal{H}^j_{\Delta\text{RS-TO}}$. If $(\mathbf{x}_o, \mathbf{y}_o) \notin \mathcal{H}^j_{\Delta\text{RS-TO}}$, program (20) is either infeasible or has a finite optimal value $\beta_o < 0$.

For each $j \in \mathcal{J}$, denote $\hat{P}^{(j)}_o$ the program obtained from program (20) by removing its constraint (20e). Let $\hat{\beta}^{(j)}_o$ be its optimal value.

**Theorem C.1.** *If program (20) has a finite (nonnegative or negative) optimal value $\beta^{(j)}_o$, then program $\hat{P}^{(j)}_o$ has a finite optimal value $\hat{\beta}^{(j)}_o$, and $\hat{\beta}^{(j)}_o = \beta_o$. If program (20) is infeasible, then either program $\hat{P}^{(j)}_o$ has a finite optimal value $\hat{\beta}^{(j)}_o < 0$ or it is infeasible.*

**Proof of Theorem C.1.** First note that the feasible region of program (20) is a subset of the feasible region of program $\hat{P}^{(j)}_o$. Therefore, we always have $\beta^{(j)}_o \leq \hat{\beta}^{(j)}_o$. (This includes the case of infeasible programs for which the optimal value is taken equal to $-\infty$.)

Assume that program (20) has a finite optimal value $\beta^{(j)}_o$. Then the value $\beta = \beta^{(j)}_o$ is achievable in program $\hat{P}^{(j)}_o$. Assume that program $\hat{P}^{(j)}_o$ has a feasible solution $\Gamma = (\delta', \boldsymbol{\pi}', \beta')$ such that $\beta' > \beta^{(j)}_o$. Because $\mathbf{y}_o, \mathbf{g}_y \geq \mathbf{0}$, we have $\mathbf{y}_o + \beta'\mathbf{g}_y \geq \mathbf{y}_o + \beta^{(j)}_o\mathbf{g}_y \geq \mathbf{0}$. Therefore, $\Gamma$ is feasible in program (20), which contradicts the assumption that $\beta^{(j)}_o \geq 0$ is its optimal value. Then $\beta \leq \beta^{(j)}_o$ in all feasible solutions of program $\hat{P}^{(j)}_o$ and $\beta^{(j)}_o$ is its optimal value.

Let $(\mathbf{x}_o, \mathbf{y}_o) \notin \mathcal{H}^j_{\Delta\text{RS-TO}}$ and let program (20) be infeasible. Assume that program $\hat{P}^{(j)}_o$ has a feasible solution with $\beta \geq 0$. Because $\mathbf{y}_o, \mathbf{g}_y \geq \mathbf{0}$, this solution also satisfies the inequality (20e). This contradicts the assumption that program (20e) is infeasible. Therefore, if program $\hat{P}^{(j)}_o$ is feasible, then $\beta < 0$ in any of its feasible solutions. Because in this case the objective function of program $\hat{P}^{(j)}_o$ is bounded above, its optimal value $\hat{\beta}^{(j)}_o$ is finite. Because $\beta < 0$ for any feasible solution of program $\hat{P}^{(j)}_o$, we also have $\hat{\beta}^{(j)}_o < 0$. □

The above theorem shows that, if $(\mathbf{x}_o, \mathbf{y}_o) \in \mathcal{H}^j_{\Delta\text{RS-TO}}$, then the optimal $\beta^{(j)}_o$ of program (20) remains unchanged if we remove constraint (20e) from the program. If $(\mathbf{x}_o, \mathbf{y}_o) \notin \mathcal{H}^j_{\Delta\text{RS-TO}}$, then program (20) and its relaxed variant without the constraint (20e) have negative optimal values or are infeasible. Neither case affects the maximum in formula (21), which is defined by the optimal values of programs (20e) for such $j$ that $(\mathbf{x}_o, \mathbf{y}_o) \in \mathcal{H}^j_{\Delta\text{RS-TO}}$. (Because $(\mathbf{x}_o, \mathbf{y}_o) \in \mathcal{T}^{\text{NC}}_{\Delta\text{RS-TO}}$, at least one such $j$ exists.) We state this result as follows.

**Corollary C.1.** *The value $\beta_o$ calculated by formula (21) remains unchanged if we solve all programs (20), $j \in \mathcal{J}$, without the constraint (20e).*

The following example illustrates Theorem C.1 in the case when $(\mathbf{x}_o, \mathbf{y}_o) \notin \mathcal{H}^j_{\Delta\text{RS-TO}}$.

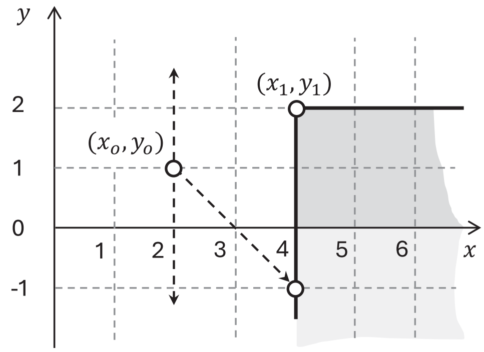


**Fig. 3.** An illustration to Example 5.

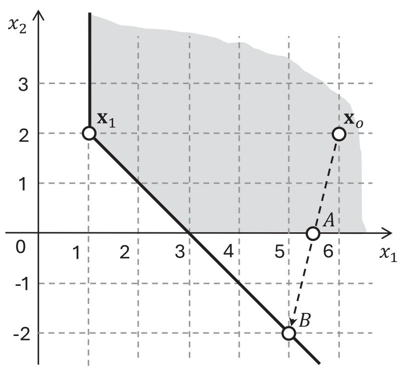


**Fig. 4.** An illustration to Example 6.

**Example 5.** For simplicity, we consider the case of VRS and assume that no trade-offs are specified. We use the simplified notation $\mathcal{H}^j_{\text{VRS}}$ for the respective hulls in this case.

Let $(x_1, y_1) = (4, 2)$ be an observed DMU, where the first component is an input and the second is an output. Fig. 3 shows the hull $\mathcal{H}^1_{\text{VRS}}$ generated by DMU $(x_1, y_1)$ without trade-offs as the dark shaded area in the nonnegative orthant. Consider the DMU $(x_o, y_o) = (2, 1)$ and the directional vector $(g_x, g_y) = (1, 1)$. It is straightforward to verify that program (20) stated for $j = 1$ (assessing the directional distance function for DMU $(x_o, y_o)$ in the hull $\mathcal{H}^1_{\text{VRS}}$) is infeasible. However, the optimal value of program (20) with constraint (20e) removed is equal to $-2$, projecting DMU $(x_o, y_o)$ on the point $(4, -1)$. If we change the directional vector $(g_x, g_y)$ to $(0, 1)$, then both program (20) and its relaxation without the constraint (20e) become infeasible.

The next example shows that, unlike the output constraint (20e), the input nonnegativity constraint (20d) is generally not redundant and should not be removed from program (20).

**Example 6.** Let $(\mathbf{x}_1, \mathbf{y}_1)$ be an observed DMU with the input and output vectors $\mathbf{x}_1 = (1, 2)^\top$ and $\mathbf{y}_1 = (1)$. Also consider the trade-off that allows us to substitute one unit of input 2 by one unit of input 1, without altering the output: $\mathbf{p} = (-1, 1)^\top$, $\mathbf{q} = (0)$.

Fig. 4 shows the section of the hull $\mathcal{H}^1_{\text{VRS-TO}}$ in the two input dimensions as the shaded area contained in the nonnegative orthant. Consider assessing the efficiency of DMU $(\mathbf{x}_o, \mathbf{y}_o)$, where $\mathbf{x}_o = (6, 2)^\top$ and $\mathbf{y}_o = (1)$, in the hull $\mathcal{H}^1_{\text{VRS-TO}}$, using the directional vector $\left(g_{x_1}, g_{x_2}, g_y\right) = (1, 4, 0)$. The optimal value of program (20) is equal to 0.5, which correctly identifies the projection of DMU $(\mathbf{x}_o, \mathbf{y}_o)$ represented by point $A$ in Fig. 4. The same program from which constraint (20d) is removed has the optimal value equal to 1 and produces an incorrect target represented by point $B$. This projection has a negative input 2 and is outside the hull $\mathcal{H}^1_{\text{VRS-TO}}$.

## References

Afsharian, M., & Podinovski, V. V. (2018). A linear programming approach to efficiency evaluation in nonconvex metatechnologies. *European Journal of Operational Research*, *268*(1), 268–280.

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.

Atici, K. B., & Podinovski, V. V. (2015). Using data envelopment analysis for the assessment of technical efficiency of units with different specialisations: An application to agriculture. *Omega*, *54*, 72–83.

Banker, R. D. (1984). Estimating most productive scale size using data envelopment analysis. *European Journal of Operational Research*, *17*(1), 35–44.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, *30*(9), 1078–1092.

Banker, R. D., & Thrall, R. M. (1992). Estimation of returns to scale using data envelopment analysis. *European Journal of Operational Research*, *62*(1), 74–84.

Bogetoft, P. (1996). DEA models on relaxed convexity assumptions. *Management Science*, *42*(3), 457–465.

Briec, W., & Kerstens, K. (2006). Input, output and graph technical efficiency measures on non-convex FDH models with various scaling laws: An integrated approach based upon implicit enumeration algorithms. *TOP*, *14*, 135–166.

Briec, W., Kerstens, K., & Vanden Eeckaut, P. (2004). Non-convex technologies and cost functions: Definitions, duality and nonparametric tests of convexity. *Journal of Economics*, *81*(2), 155–192.

Chambers, R. G., Chung, Y., & Färe, R. (1998). Profit, directional distance functions, and nerlovian efficiency. *Journal of Optimization Theory and Applications*, *98*(2), 351–364.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429–444.

Cherchye, L., Kuosmanen, T., & Post, T. (2000). What is the economic meaning of FDH? A reply to Thrall. *Journal of Productivity Analysis*, *13*(3), 259–263.

Cherchye, L., Kuosmanen, T., & Post, T. (2001). FDH directional distance functions: With an application to European commercial banks. *Journal of Productivity Analysis*, *15*(3), 201–215.

Deprins, D., Simar, L., & Tulkens, H. (1984). Measuring labor-efficiency in post offices. In M. Marchand, P. Pestieau, & H. Tulkens (Eds.), *The performance of public enterprises: Concepts and measurements* (pp. 243–267). Amsterdam: North-Holland.

Färe, R., Grosskopf, S., & Lovell, C. A. K. (1983). The structure of technical efficiency. *Scandinavian Journal of Economics*, *85*(2), 181–190.

Färe, R., Grosskopf, S., & Lovell, C. A. K. (1985). *The measurement of efficiency of production*. Boston: Kluwer Academic Publishers.

Førsund, F. R., Hjalmarsson, L., Krivonozhko, V. E., & Utkin, O. B. (2007). Calculation of scale elasticities in DEA models: Direct and indirect approaches. *Journal of Productivity Analysis*, *28*(1), 45–56.

Hadjicostas, P., & Soteriou, A. C. (2006). One-sided elasticities and technical efficiency in multi-output production: A theoretical framework. *European Journal of Operational Research*, *168*(2), 425–449.

Kerstens, K., Sadeghi, J., Toloo, M., & Van de Woestyne, V. (2022). Procedures for ranking technical and cost efficient units: With a focus on nonconvexity. *European Journal of Operational Research*, *300*(1), 269–281.

Kerstens, K., & Vanden Eeckaut, P. (1999). Estimating returns to scale using nonparametric deterministic technologies: A new method based on goodness-of-fit. *European Journal of Operational Research*, *113*(1), 206–214.

Kerstens, K., & Van de Woestyne, I. (2021). Cost functions are nonconvex in the outputs when the technology is nonconvex: Convexification is not harmless. *Annals of Operations Research*, *305*(1–2), 81–106.

Koopmans, T. C. (1957). *Three essays on the state of economic science*. New York: McGraw Hill.

Mehdiloo, M., Sadeghi, J., & Kerstens, K. (2024). *Top down axiomatic modeling of metatechnologies and evaluating directional economic efficiency*: *IESEG working paper 2024-EQM-03*, Lille: IESEG School of Management.

Papaioannou, G., & Podinovski, V. V. (2023). Multicomponent production technologies with restricted allocations of shared inputs and outputs. *European Journal of Operational Research*, *308*(1), 274–289.

Papaioannou, G., & Podinovski, V. V. (2024). Free disposal hull models of multicomponent technologies. *Annals of Operations Research*, http://dx.doi.org/10.1007/s10479-024-06140-z.

Podinovski, V. V. (2004a). Efficiency and global scale characteristics on the no free lunch assumption only. *Journal of Productivity Analysis*, *22*(3), 227–257.

Podinovski, V. V. (2004b). Local and global returns to scale in performance measurement. *Journal of the Operational Research Society*, *55*(2), 170–178.

Podinovski, V. V. (2004c). On the linearisation of reference technologies for testing returns to scale in FDH models. *European Journal of Operational Research*, *152*(3), 800–802.

Podinovski, V. V. (2004d). Production trade-offs and weight restrictions in data envelopment analysis. *Journal of the Operational Research Society*, *55*(12), 1311–1322.

Podinovski, V. V. (2017). Returns to scale in convex production technologies. *European Journal of Operational Research*, *258*(3), 970–982.

Podinovski, V. V., & Bouzdine-Chameeva, T. (2013). Weight restrictions and free production in data envelopment analysis. *Operations Research*, *61*(2), 426–437.

Podinovski, V. V., & Bouzdine-Chameeva, T. (2015). Consistent weight restrictions in data envelopment analysis. *European Journal of Operational Research*, *244*(1), 201–209.

Podinovski, V. V., Chambers, R. G., Atici, K. B., & Deineko, I. D. (2016). Marginal values and returns to scale for nonparametric production frontiers. *Operations Research*, *64*(1), 236–250.

Podinovski, V. V., Wu, J., & Argyris, N. (2024). Production trade-offs in models of data envelopment analysis with ratio inputs and outputs: An application to schools in England. *European Journal of Operational Research*, *313*(1), 359–372.

Ramón, N., Ruiz, J. L., & Sirvent, I. (2016). On the use of DEA models with weight restrictions for benchmarking and target setting. In J. Aparicio, C. A. K. Lovell, & J. T. Pastor (Eds.), *Advances in efficiency and productivity* (pp. 149–180). New York: Springer Science Business Media.

Sahoo, B. K., & Tone, K. (2015). Scale elasticity in non-parametric DEA approach. In J. Zhu (Ed.), *Data envelopment analysis: A handbook of models and methods* (pp. 269–290). New York: Springer Science Business Media.

Scarf, H. E. (1981). Production sets with indivisibilities. Part I: Generalities. *Econometrica*, *49*(1), 1–32.

Tavakoli, I. M., & Mostafaee, A. (2019). Free disposal hull efficiency scores of units with network structures. *European Journal of Operational Research*, *277*(3), 1027–1036.

Thanassoulis, E., Kortelainen, M., Johnes, G., & Johnes, J. (2011). Costs and efficiency of higher education institutions in England: A DEA analysis. *Journal of the Operational Research Society*, *62*(7), 1282–1297.

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.

Tone, K., & Sahoo, B. K. (2003). Scale, indivisibilities and production function in data envelopment analysis. *International Journal of Production Economics*, *84*(2), 165–192.

Tulkens, H. (1993). On FDH efficiency: Some methodological issues and application to retail banking, courts, and urban transit. *Journal of Productivity Analysis*, *4*(1), 183–210.

Tulkens, H., & Vanden Eeckaut, P. (1995). Non-parametric efficiency, progress and regress measures for panel data: Methodological aspects. *European Journal of Operational Research*, *80*(3), 474–499.