City Research Online

City, University of London Institutional Repository

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

A unified approach to radial, hyperbolic, and directional efficiency measurement in Data Envelopment Analysis

Margaréta Halická[a], Mária Trnovská[a,*], Aleš Černý[b]

[a]*Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava, Mlynská dolina, 842 48 Bratislava, Slovakia*
[b]*Bayes Business School, City, University of London, 106 Bunhill Row, London EC1Y 8TZ, UK*

Abstract

The paper analyses properties of a large class of "path-based" Data Envelopment Analysis models through a unifying general scheme. The scheme includes the well-known oriented radial models, the hyperbolic distance function model, the directional distance function models, and even permits their generalisations. The modelling is not constrained to non-negative data and is flexible enough to accommodate variants of standard models over *arbitrary data*.

Mathematical tools developed in the paper allow systematic analysis of the models from the point of view of ten desirable properties. It is shown that some of the properties are satisfied (resp., fail) for all models in the general scheme, while others have a more nuanced behaviour and must be assessed individually in each model. Our results can help researchers and practitioners navigate among the different models and apply the models to mixed data.

Keywords: Data envelopment analysis, Radial efficiency measures, Hyperbolic distance function, Directional distance function, Negative data

1. Introduction

Data Envelopment Analysis (DEA) is a non-parametric analytical method used to assess the performance of Decision Making Units (DMU). The DEA approach defines technology sets via observed input-output data in combination with certain axioms. Each traditional model links an efficiency measure and a given technology set using a mathematical optimisation programme. By solving the programme for an assessed DMU, one finds both the efficiency score and a projection[1] on the frontier of the technology set that dominates the assessed DMU.[2] Russell and Schworm (2018) distinguish between two main methods of arriving at the efficiency score, giving rise to two classes of DEA models appositely named *slacks-based* and *path-based* models.

---

*Corresponding author

Email addresses: `halicka@fmph.uniba.sk` (Margaréta Halická), `trnovska@fmph.uniba.sk` (Mária Trnovská), `ales.cerny.1@city.ac.uk` (Aleš Černý)

[1]Synonymously with "projection," the literature also uses the terminology "reference unit," "projection benchmark," "projection point," or "target" in this context.

[2]The resulting projection need not be Pareto–Coopmans efficient.

This paper establishes a *general path-based scheme* built on variable returns to scale (VRS) technology that provides a unifying mathematical framework for the well-known oriented radial models, the hyperbolic distance function model, and the directional distance function models. The scheme offers a rich menu of projection paths for efficiency analysis and the special capability of handling negative data, which is important in a variety of applications, especially in the areas of accounting and finance. In the confines of the general scheme, we provide a systematic and comprehensive analysis of the models *vis-à-vis* ten desirable properties. Overall, our paper amplifies the message of Russell and Schworm (2018) that recognising the class to which a model belongs allows one to deduce some of the model characteristics, which is important for the correct interpretation of models and their practical use.

The slacks-based and the path-based models are easily distinguishable from each other on the basis of the objective function in the corresponding programmes, or, in other words, based on the way they look for a projection. According to Russell and Schworm (2018), the models in the *slacks-based class* search for the projection by "specifying the form of aggregation over the coordinate-wise slacks." The slacks indicate the input surplus and output shortage between the projection and the assessed DMU. The main representatives of this class are the Slacks-Based Measure (SBM) model of Tone (2001), the Russell Graph Measure model of Färe, Grosskopf and Lovell (1985), the Additive Model (AM) of Charnes et al. (1985), and the Weighted Additive Models (WAM) including the Range Adjusted Measure (RAM) model (Cooper, Park and Pastor, 1999) and the Bounded Adjusted Measure (BAM) model (Cooper et al., 2011).

The *path-based models* — the main focus of this paper — search for the projection by specifying various parametric paths which run from the assessed DMU to the boundary of the technology set. In the special case of the radial BCC input or output-oriented models (Banker, Charnes and Cooper, 1984), the path is defined by a ray connecting the DMU to the origin in the space of inputs or outputs and thus represents a proportional, radial contraction of inputs or expansion of outputs. In the Directional Distance Function (DDF) model (Chambers, Chung and Färe, 1996a, 1998), the path is determined by a ray in a pre-assigned direction pointing from the assessed DMU towards the dominating part of the frontier.[3] On this path, one then seeks the point of minimal distance to the frontier of the technology set. The Hyperbolic Distance Function (HDF) model (Färe, Grosskopf and Lovell, 1985) combines together the input and output oriented BCC models by using a hyperbolic path that allows for a simultaneous equiproportionate contraction of inputs and expansion of outputs.

Despite the many papers on DEA, there are only a few studies analysing the properties of either class of models in a unified framework. In the context of general economic productivity theory, a series of articles Russell and Schworm (2008), Russell and Schworm (2011), Levkoff, Russell and Schworm (2011), Roshdi, Hasannasab, Margaritis and Rouse (2018), and Hasannasab,

---

[3]There are also approaches, where the directional vector may not point to the dominating part of the frontier, or may even be determined endogenously. For a discussion on these non-standard approaches we refer to Pastor et al. (2022).

Margaritis, Roshdi and Rouse (2019) provided a comprehensive analysis of efficiency measures over different types of productivity sets. The main message from these papers is that the slacks-based measures, when operating in the full input-output space, identify the Pareto–Koopmans efficiency unambiguously while the path-based measures do not.

In the DEA setting, Cooper, Park and Pastor (1999) and Sueyoshi and Sekitani (2009) analyse efficiency measures and introduce a set of desirable properties that an ideal DEA model should satisfy. These properties include the indication of strong efficiency; boundedness; strict monotonicity; unit invariance; and translation invariance. The selected DEA models are then classified on the basis of these criteria.

In other work, Halická and Trnovská (2021) analyse the properties of slacks-based models in a general scheme that encompasses all commonly used models in this class, but also allows for the construction of new models.[4] All models in the general slacks-based scheme project onto the strongly efficient frontier and therefore account for all sources of inefficiencies. Among them, the RAM model performs best when measured against eight desirable properties, satisfying seven, and failing only the unique projection property due to its linearity. Recently, Aparicio and Monge (2022) have proposed a convex generalisation of the RAM model that falls into the general scheme of Halická and Trnovská (2021) and provides a unique projection point. Thus, the new model currently claims the top spot in terms of the number of desirable properties.

Our paper aims to redress the lack of comprehensive analysis of path-based models in the literature. The knowledge about the path-based models is currently fragmented across many articles with varying focus. This situation is exacerbated by the fact that the properties of, for example, DDF models depend significantly on the choice of direction vectors and these have not been treated systematically to date. There appears to be a general consensus that the path-based models do not guarantee strongly efficient projection points and, therefore, their efficiency score is overstated; and that they are monotone but not strictly monotone. Many authors noticed difficulties with super-efficiency measurement under variable returns to scale and the associated measurement of productivity change over time (e.g., Briec and Kerstens, 2009). Aparicio et al. (2016) investigated the translation invariance of DDF models. A certain type of homogeneity was observed in the oriented radial models and the HDF model (Cuesta and Zofío, 2005). On the other hand, DDF models have the property of homogeneity only in the case of constant returns to scale. Another type of homogeneity (so-called g-homogeneity) was introduced to describe some properties of DDF (Hudgins and Primont, 2007).

In this paper, we analyse path-based models in a general framework. The main building block of our approach is a parametric path that starts at the assessed DMU and for decreasing values of the parameter passes through dominating units in the technology set towards its frontier. The path has two main ingredients: (i) a direction vector; (ii) a real-valued smooth function, whose

---

[4]The terminology "non-radial" and "radial pattern models" was used for slacks-based and path-based models, respectively.

specifications determine the path shape and, ultimately, the model properties. The assumptions imposed on directions and path functions are flexible enough to accommodate all standard models, such as the BCC input and output-oriented models, the HDF and DDF models, the generalised distance function model by Chavas and Cox (1999), and even offer the possibility of going beyond.

A further advantage of the proposed general framework is that it permits extension of existing models designed for non-negative data to arbitrary data (i.e., it accommodates the input and output data, for which some or all components are negative). As a rule, DEA models are designed for non-negative data. In practice, however, negative data are encountered in many applications in areas such as insurance, accounting, finance, or banking. A common approach to overcoming this difficulty is to use translation-invariant models, which can be applied to arbitrary data without modifications. The drawback is that only a few standard models have the property of translation invariance. These include some of the DDF models (see Aparicio et al., 2016) but not BCC or HDF. Therefore, many DEA studies propose procedures of varying complexity to deal with negative data (e.g., Cheng et al., 2013; Tone et al., 2020).

The general framework for path-based models allows us to survey in one place the properties of all standard path-based models and also offer certain guidance on how to construct new models with given properties. The desirable properties analysed in this paper include (a) unique projection point; (b) indication; (c) strong efficiency of projection points; (d) boundedness; (e) unit and translation invariance; (f) (strict) monotonicity; (g) super-efficiency; and (h) homogeneity. Our results divide the properties into two groups: the properties that hold universally (resp., universally fail) for every model in the general framework; and the properties that must be assessed individually for each model. With the help of the general framework, we show that only the unique projection point property is satisfied in all models. Three other properties (indication, strong efficiency of projection, and strict monotonicity) usually fail simultaneously, although surprisingly there are very special cases where they simultaneously hold. For properties in the second group, the article provides tools to determine whether a property holds or fails in a specific model. In particular, the monotonicity property is satisfied by all standard path-based models, but the homogeneity property is limited to only a specific type of directions and path functions. We find that there are no trade-offs between the properties of indication, strong efficiency, and strict monotonicity on the one hand and the property of homogeneity on the other.

The paper is organised as follows. Section 2 introduces basic terminology and notation concerning, among others, the technological set and its efficient frontier over general data; desirable properties of DEA models; and standard path-based models. Section 3 proposes a general scheme for path-based models, analyses its basic properties, and discusses its geometry. Section 4 conducts a deeper analysis of the general scheme in light of ten desirable properties. This section develops practical criteria for each property and illustrates them on individual standard path-based models. Section 5 extends the above analysis in two directions: (a) properties of standard models over arbitrary data; (b) construction of new models with good properties. Section 6 concludes.

2. Preliminaries

Let us establish some notation. $\mathbb{R}^d$ denotes the d dimensional Euclidean space, and $\mathbb{R}^d_+$ its non-negative orthant. The lowercase bold letters denote column vectors and the uppercase bold letters matrices. The superscript T denotes the transpose of a column vector or a matrix. The symbol e denotes a vector of ones.

Consider a set of n observed decision-making units DMU_j $(j = 1, \ldots, n)$, each consuming m inputs x_{ij} $(i = 1, \ldots, m)$ to produce s outputs y_{rj} $(r = 1, \ldots, s)$. For each $j = 1, \ldots, n$, the data of the inputs and outputs of DMU_j can be arranged into the column vectors $\boldsymbol{x}_j = (x_{1j}, \ldots, x_{mj})^T \in \mathbb{R}^m$ of the inputs and $\boldsymbol{y}_j = (y_{1j}, \ldots, y_{sj})^T \in \mathbb{R}^s$ of the outputs. Finally, the input and output vectors of all DMUs form the $m \times n$ input and $s \times n$ output matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, i.e., $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]$ and $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n]$, respectively.

In the article, we reserve the notations i, j, and r for the indices that go through whole index sets $\{1, \ldots, m\}$, $\{1, \ldots, n\}$, and $\{1, \ldots, s\}$, respectively. We make no assumptions about the non-negativity of the data at this point. The non-negativity requirement may follow later from other assumptions, and we shall alert the reader whenever that is the case.

2.1. Technology set

On the basis of the given data, we consider the following technology set:

$$\mathcal{T} = \left\{ (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^m \times \mathbb{R}^s \mid \boldsymbol{X}\boldsymbol{\lambda} \leq \boldsymbol{x}, \ \boldsymbol{Y}\boldsymbol{\lambda} \geq \boldsymbol{y}, \ \boldsymbol{\lambda} \geq \boldsymbol{0}, \ \boldsymbol{e}^T\boldsymbol{\lambda} = 1 \right\}, \tag{1}$$

corresponding to variable returns to scale (VRS). Note that the common non-negativity of $(\boldsymbol{x}, \boldsymbol{y})$ is not imposed here. Elements of $\mathcal{T}$ will be called units. It follows from (1) that the closed set $\mathcal{T}$ has a non-empty interior; we shall denote its boundary by $\partial\mathcal{T} := \mathcal{T} \setminus \mathrm{int}\mathcal{T}$.

By $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ we denote a unit from $\mathcal{T}$ to be currently evaluated. For input vectors, we also use the notation $\boldsymbol{x}^{\min}, \boldsymbol{x}^{\max}, \boldsymbol{x}^{ev}, \boldsymbol{x}^{sd}$, where for $i = 1, \ldots m$ we set $x_i^{\min} = \min_j x_{ij}$, $x_i^{\max} = \max_j x_{ij}$, $x_i^{ev} = \frac{1}{n}\sum_j x_{ij}$, and $x_i^{sd} = \sqrt{\frac{1}{n}\sum_j (x_{ij} - \boldsymbol{x}_i^{ev})^2}$. Here x_i^{sd} is the standard deviation of the i-th input over all DMU_j, $j = 1, \ldots, n$. The notation $\boldsymbol{y}^{\min}$, $\boldsymbol{y}^{\max}$, $\boldsymbol{y}^{ev}$, and $\boldsymbol{y}^{sd}$ is introduced analogously for the output vectors. Without loss of generality, we assume that $\boldsymbol{x}^{\min} < \boldsymbol{x}^{\max}$ and $\boldsymbol{y}^{\min} < \boldsymbol{y}^{\max}$. Otherwise, there would be components of inputs / outputs, where all DMUs take the same value, and such components can be excluded from the analysis.

We write $(\boldsymbol{x}, \boldsymbol{y}) \succsim (\boldsymbol{x}_o, \boldsymbol{y}_o)$ if the unit $(\boldsymbol{x}, \boldsymbol{y})$ *dominates* the unit $(\boldsymbol{x}_o, \boldsymbol{y}_o)$, that is, if $\boldsymbol{x} \leq \boldsymbol{x}_o$ and $\boldsymbol{y} \geq \boldsymbol{y}_o$. A unit $(\boldsymbol{x}, \boldsymbol{y})$ *strictly dominates* the unit $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ if $\boldsymbol{x} < \boldsymbol{x}_o$ and $\boldsymbol{y} > \boldsymbol{y}_o$. A unit $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$ is called *strongly efficient* if there is no other unit in $\mathcal{T}$ that dominates $(\boldsymbol{x}_o, \boldsymbol{y}_o)$, that is, if $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{T}$ dominates $(\boldsymbol{x}_o, \boldsymbol{y}_o)$, then $(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}_o, \boldsymbol{y}_o)$.[5] A unit $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$ is called *weakly efficient* if there is no unit in $\mathcal{T}$ that strictly dominates $(\boldsymbol{x}_o, \boldsymbol{y}_o)$. Evidently, any strongly efficient unit is weakly efficient, and weakly efficient units lie on the boundary $\partial\mathcal{T}$.

---

[5]This is the well known Pareto–Koopmans efficiency. Some authors call such units Pareto efficient, or fully efficient; see the discussion in Cooper et al. (2007, p. 45).

The converse is also true: every unit on the boundary $\partial\mathcal{T}$ is weakly efficient because the definition of $\mathcal{T}$ in (1) does not impose the non-negativity assumption on the units therein. Therefore, the boundary $\partial\mathcal{T}$ is partitioned into the *strongly efficient frontier* $\partial^S\mathcal{T}$ containing all strongly efficient units and the remaining part $\partial^W\mathcal{T} := \partial\mathcal{T} \setminus \partial^S\mathcal{T}$, which consists of weakly but not strongly efficient units. In this paper, we refer to the remaining part of the boundary as *weakly efficient frontier*. One thus has $\partial\mathcal{T} = \partial^S\mathcal{T} \cup \partial^W\mathcal{T}$ and $\partial^S\mathcal{T} \cap \partial^W\mathcal{T} = \emptyset$. The simple proof of the next lemma is omitted.

Lemma 1. *For $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$, the following statements hold.*

(a) *$(\boldsymbol{x}^{\min}, \boldsymbol{y}^{\max}) \succsim (\boldsymbol{x}_o, \boldsymbol{y}_o)$;*

(b) *$(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \partial\mathcal{T}$ if and only if for all $(\boldsymbol{d}^x, \boldsymbol{d}^y) \geq \boldsymbol{0}$ such that $(\boldsymbol{x}_o - \boldsymbol{d}^x, \boldsymbol{y}_o + \boldsymbol{d}^y) \in \mathcal{T}$, one has $(\boldsymbol{d}^x, \boldsymbol{d}^y) \not> \boldsymbol{0}$;*

(c) *$(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \partial^S\mathcal{T}$ if and only if for all $(\boldsymbol{d}^x, \boldsymbol{d}^y) \geq \boldsymbol{0}$ such that $(\boldsymbol{x}_o - \boldsymbol{d}^x, \boldsymbol{y}_o + \boldsymbol{d}^y) \in \mathcal{T}$, one has $(\boldsymbol{d}^x, \boldsymbol{d}^y) = \boldsymbol{0}$;*

(d) *$(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \partial^W\mathcal{T}$ if and only if there exists $(\boldsymbol{d}^x, \boldsymbol{d}^y) \gneq \boldsymbol{0}$ such that $(\boldsymbol{x}_o - \boldsymbol{d}^x, \boldsymbol{y}_o + \boldsymbol{d}^y) \in \mathcal{T}$ and if for all $(\boldsymbol{d}^x, \boldsymbol{d}^y) \geq \boldsymbol{0}$ such that $(\boldsymbol{x}_o - \boldsymbol{d}^x, \boldsymbol{y}_o + \boldsymbol{d}^y) \in \mathcal{T}$, one has $(\boldsymbol{d}^x, \boldsymbol{d}^y) \not> \boldsymbol{0}$.*

2.2. Efficiency measures and their desirable properties

On the technology set $\mathcal{T}$, one can define various *measures*, which then give rise to a specific DEA *models*. This is achieved by formulating certain mathematical programming problems that are applied separately to each evaluated unit $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$.

A mathematical programme consists of an *objective function* and a set of conditions for *decision variables*. The values of the decision variables that satisfy these conditions are called *feasible solutions* of the programme. Solving the programme yields the *optimal value* of the objective function and a subset of feasible solutions called *optimal solutions*. The optimal value of a DEA model determines the *efficiency score / value of the measure* of the evaluated unit. Optimal solutions also yield virtual units called *projections* of the evaluated unit — the elements of $\partial\mathcal{T}$ that dominate $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ and are associated with the efficiency score.

Depending on the chosen measure, models and their efficiency scores have varying properties. Following the work of Färe and Lovell (1978), Pastor et al. (1999), Sueyoshi and Sekitani (2009), Russell and Schworm (2018), Aparicio and Monge (2022), and Halická and Trnovská (2021), we highlight ten desirable properties for efficiency measurement in DEA.

(P1) **Unique projection for efficiency comparison.** While the model may admit multiple optimal solutions, the projection point is unique.

(P2) **Indication.** The value of the measure equals one if and only if the evaluated unit is strongly efficient.

(P3) **Strong efficiency of projections.** Projections generated by the measure are strongly efficient.

(P4) **Boundedness.** The measure takes values between zero and one.

(P5) **Units invariance.** The value of the measure does not depend on the units of measurement in the input and output variables.

(P6) **Translation invariance.** The value of the measure is not affected by translation of inputs or outputs.

(P7) **Monotonicity.** An increase in any input or a decrease in any output relative to the evaluated unit, keeping other inputs and outputs constant, reduces or maintains the value of the measure.

(P8) **Strict monotonicity.** An increase in any input or a decrease in any output relative to the evaluated unit, keeping other inputs and outputs constant, reduces the value of the measure.

(P9) **Super-efficiency.** The value of the measure of a unit outside the technology set is well defined and finite.

(P10) **Homogeneity.** Feasible scaling of the input vector and the output vector of the evaluated unit by a power of $\mu > 0$ (the power may be different for inputs and outputs) results in the efficiency score being scaled by a power of μ.

The importance of individual properties varies with the goals of the analysis. In the first instance, one needs to know whether the efficiency score equal to one indicates the strong efficiency of the evaluated unit. Therefore, (P2) is perhaps the most important property. It is closely related to (P3), although the two properties are not equivalent. Failure of (P3) indicates that the score may be overestimated. Some models are ambiguous in terms of satisfying the properties (P2) and (P3) and require further analysis to reach a conclusion.

Invariance properties (P5) and (P6) are important when the input and output data are expressed in different measurement units or when the data also contain negative values, respectively. No universal conclusions can be drawn regarding these properties, hence each DEA model must be assessed individually. The property (P6) is tied to variable returns-to-scale technology set.

The monotonicity properties (P7) and (P8) ensure that competing units are evaluated with a certain degree of fairness, in that a dominating unit receives a score not below the score of a dominated unit. The weak monotonicity (P7) is usually satisfied in standard DEA models and the strong monotonicity (P8) is satisfied in most slacks-based models.

The property (P9) is important for (i) measuring super-efficiency, which allows one to discriminate among efficient units; and (ii) measuring the frontier shift between different time periods; see Sueyoshi and Sekitani (2009). Considerable difficulties with this property relate particularly to VRS, where some super-efficiency models are generally not feasible.

The homogeneity property (P10), together with (P2), (P3), and (P8), was formulated for the first time as a desirable property of efficiency measures in Färe and Lovell (1978) in connection with input-oriented measures. Since then, it has been subject to several generalisations suitable for output-oriented models and graph models, which is also reflected in the terminology (e.g., "almost homogeneity"). For slacks-based models, the concept was relaxed to sub-homogeneity (Pastor et al., 1999), and for the directional distance function, it was modified to g-homogeneity (Hudgins and Primont, 2007).

The property of boundedness (P4) was included in the list of desirable properties in Färe et al. (1983). The property (P4) guarantees that the efficiency measure is bounded by 0 and 1, where the value 1 identifies the strong efficiency in (P2).

In contrast to all the remaining properties that are expressed in terms of the efficiency measure, the properties (P1) and (P3) are expressed using the projections provided by the models. The importance of (P1) was highlighted in Sueyoshi and Sekitani (2009), especially in connection with the slacks-based models, where the concept of projection is not uniquely defined.

It is well known that no model meets all properties (P1)–(P10). However, some of the properties occur in groups and hold simultaneously for a class of models. Halická and Trnovská (2021) show that slacks-based models satisfy (P2), (P3), (P7), (P8), and violate (P1). The present paper shows that path-based models satisfy (P1) and (P7), and fail (P2), (P3), and (P8).

We shall now introduce some standard models that enter our analysis as special cases. We formulate these models in a way that allows for certain generalisations.

2.3. Standard path-based models

First, we present input/output-oriented radial models, which are known in DEA under the BBC acronym in honour of their authors Banker et al. (1984), who introduced them to DEA in the VRS version. Note that the standard input/output-oriented DEA radial models are inspired by the Shephard's (Shephard, 1953) input/output distance functions, and the technical efficiency measurement is based on the seminal work of Farrell (1957). The standard DEA formulation of these models is over non-negative data with the assumption

$$x_j \gneq 0 \quad \text{and} \quad y_j \gneq 0 \qquad \text{for each } j \in \{1, \ldots, n\}. \tag{2}$$

The input oriented BCC model can be formulated as

$$\begin{aligned} \text{BCC-I:} \quad \min \quad & \theta \\ & X\lambda \leq \theta x_o \\ & Y\lambda \geq y_o \\ & e^T\lambda = 1, \ \lambda \geq 0, \end{aligned} \tag{3}$$

and the output oriented BCC model as

$$\begin{aligned} \text{BCC-O:} \quad \min \quad & \theta \\ & X\lambda \leq x_o \\ & Y\lambda \geq \frac{1}{\theta}y_o \\ & e^T\lambda = 1, \ \lambda \geq 0. \end{aligned} \tag{4}$$

If we denote $\psi := \frac{1}{\theta}$ and replace $\min \theta$ by $\max \psi$, then we obtain the output BCC model in standard form, where the optimal $\psi^* \geq 1$ and the efficiency score is $\frac{1}{\psi^*}$. Among the properties (P1)–(P9), the BCC models satisfy only (P1), (P4), (P5), and (P7).

The hyperbolic distance function model introduced by Färe et al. (1985) combines the input and output oriented radial measures into one measure. Commonly, it is defined over positive data only, i.e.

$$\boldsymbol{x}_j > \boldsymbol{0} \quad \text{and} \quad \boldsymbol{y}_j > \boldsymbol{0} \qquad \text{for each } j \in \{1, \dots, n\}. \tag{5}$$

In our formulation, it reads:

$$
\begin{aligned}
\text{HDF:} \qquad \min \quad & \theta \\
& \boldsymbol{X\lambda} \leq \theta \boldsymbol{x}_o \\
& \boldsymbol{Y\lambda} \geq \tfrac{1}{\theta} \boldsymbol{y}_o \\
& \boldsymbol{e}^T \boldsymbol{\lambda} = 1, \; \boldsymbol{\lambda} \geq \boldsymbol{0}.
\end{aligned} \tag{6}
$$

Halická and Trnovská (2019) investigated the properties of this model, proposed computational methods for its solution, and derived its dual form through SDP, while Hasannasab et al. (2019) established duality through the second-order cone. HDF satisfies the properties (P1), (P4), (P5), (P7), and (P9).

The use of linearisation $\frac{1}{\theta} - 1 \sim 1 - \theta$ around $\theta = 1$ in the HDF model (see Färe et al., 2016) leads to the following:

$$
\begin{aligned}
\text{DDF:} \qquad \min \quad & \theta \\
& \boldsymbol{X\lambda} \leq \boldsymbol{x}_o - (1 - \theta)\boldsymbol{x}_o \\
& \boldsymbol{Y\lambda} \geq \boldsymbol{y}_o + (1 - \theta)\boldsymbol{y}_o \\
& \boldsymbol{e}^T \boldsymbol{\lambda} = 1, \; \boldsymbol{\lambda} \geq \boldsymbol{0}.
\end{aligned} \tag{7}
$$

If we denote $\delta = 1 - \theta$ and replace $\min \theta$ by $\max \delta$, we obtain a special version of the directional distance function model with the directional vector $(g_o^x, g_o^y) = (x_o, y_o)$. This model, introduced by Briec (1997), is known as the proportional directional distance function. It satisfies the properties (P1), (P4), (P5), (P7), and (P9).

Now, we present the general directional distance function model (DDF-g). Let $\boldsymbol{g}_o^x \in \mathbb{R}_+^m$, $\boldsymbol{g}_o^y \in \mathbb{R}_+^s$ be given directional vectors that may depend on $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ and such that at least one component of $\boldsymbol{g}_o^x$ or $\boldsymbol{g}_o^y$ is positive, that is,

$$(\boldsymbol{g}_o^x, \boldsymbol{g}_o^y) \gneqq \boldsymbol{0}. \tag{8}$$

The model admits negative data, provided that the assumption of non-negativity (8) on the directional vectors is satisfied. The model reads as follows:

$$
\begin{aligned}
\text{DDF-g:} \qquad \min \quad & \theta \\
& \boldsymbol{X\lambda} \leq \boldsymbol{x}_o - (1 - \theta)\boldsymbol{g}_o^x \\
& \boldsymbol{Y\lambda} \geq \boldsymbol{y}_o + (1 - \theta)\boldsymbol{g}_o^y \\
& \boldsymbol{e}^T \boldsymbol{\lambda} = 1, \; \boldsymbol{\lambda} \geq \boldsymbol{0}.
\end{aligned} \tag{9}
$$

Once again, if we denote $\delta = 1 - \theta$ and replace $\min \theta$ by $\max \delta$, we obtain the standard form of DDF-g introduced in Chambers et al. (1996a, 1998).

Notation	$\boldsymbol{g}_o^x$	$\boldsymbol{g}_o^y$	Reference
(G1)	$\boldsymbol{x}_o$	$\boldsymbol{y}_o$	Chambers et al. (1996b)
(G2)	$\boldsymbol{x}_o - \boldsymbol{x}^{\min}$	$\boldsymbol{y}^{\max} - \boldsymbol{y}_o$	Portela et al. (2004)
(G3)	$\boldsymbol{x}^{\max} - \boldsymbol{x}^{\min}$	$\boldsymbol{y}^{\max} - \boldsymbol{y}^{\min}$	Portela et al. (2004)
(G4)	$\boldsymbol{x}^{ev}$	$\boldsymbol{y}^{ev}$	Aparicio et al. (2013)
(G5)	$\boldsymbol{x}^{sd}$	$\boldsymbol{y}^{sd}$	
(G6)	$\boldsymbol{e} = (1, \ldots, 1)^T$	$\boldsymbol{e} = (1, \ldots, 1)^T$	Chambers et al. (1996b)

Table 1: Commonly encountered choices of directions g_o. For the notation $\boldsymbol{x}^{\min}$, $\boldsymbol{x}^{\max}$, etc., see Subsection 2.1.

Since DDF is a linearisation of HDF, it is natural to ask whether there exists a generalisation of HDF whose linearisation corresponds to DDF-g. An affirmative answer is provided by the following model, which we refer to as the general hyperbolic distance function (HDF-g) model:

$$
\begin{aligned}
\text{HDF-g:} \qquad \min \quad & \theta \\
& \boldsymbol{X}\boldsymbol{\lambda} \leq \boldsymbol{x}_o - (1-\theta)\boldsymbol{g}_o^x \\
& \boldsymbol{Y}\boldsymbol{\lambda} \geq \boldsymbol{y}_o + (\tfrac{1}{\theta}-1)\boldsymbol{g}_o^y \\
& \boldsymbol{e}^T\boldsymbol{\lambda} = 1, \ \boldsymbol{\lambda} \geq \boldsymbol{0}.
\end{aligned}
\tag{10}
$$

This new model, too, uses pre-specified directions and admits negative data as long as the directions satisfy the non-negativity condition (8). It is easy to see that the choices $\boldsymbol{g}_o^x = \boldsymbol{x}_o$ and $\boldsymbol{g}_o^y = \boldsymbol{y}_o$ lead to the HDF model.

Without specific knowledge of the directional vector $\boldsymbol{g}$ beyond (8), only the property (P1) is ensured in HDF-g and DDF-g. More must be said about the choice of directional vectors to ensure the validity of the remaining properties (P2)–(P9). For a discussion of the available choices of directional vectors, see, for example, Färe et al. (2008) and Pastor et al. (2022). Table 1 presents the directions $(\boldsymbol{g}_o^x, \boldsymbol{g}_o^y)$ that will be included in the subsequent analysis of the families of HDF-g and DDF-g models. These directions exhibit a range of properties, as documented in Table 3.

(G1) is perhaps the most widely chosen direction in the family of DDF-g models, which yields the Farrell proportional distance function (see Briec and Kerstens, 2009). The directions (G2) and (G3) appear in the range-directional models developed by Portela et al. (2004). The use of (G4) directions is discussed in Aparicio et al. (2013). It is also quite common to consider DDF-g with (G6) directions, which is mathematically equivalent to minimising the l_∞ distance to the boundary of $\mathcal{T}$ (see Briec and Lesourd, 1999).

Remark 1. It is interesting to note that the expressions for directions (G1)–(G6) appear in the denominator of weights in specific WAM models. Indeed, direction (G6) is linked to the weights in AD model; directions (G2) and (G3) are linked to the weights in BAM and RAM models, respectively; direction (G1) corresponds to the weights in so-called Measure of Inefficiency Proportions (MIP) model by Cooper et al. (1999); and the weights generated via (G5) appear in the WAM model of Ševčovič et al. (2001).

3. A general scheme for path-based models

On comparing (9) and (10), we observe that the DDF-g and HDF-g models are qualitatively similar, differing only in the nature of their output bounds, which are linear for DDF-g but nonlinear for HDF-g. The latter is formalised through the function $\psi^y(\theta) = \frac{1}{\theta}$.

We shall now generalise this approach in the direction of *convex* functions ψ^y on the output side and *concave* functions ψ^x on the input side. Our new model will encompass all the models mentioned up to this point, allowing for a unified analysis of their properties. At the same time, generalisation will allow the construction of new models whose properties can be specified in advance.

By the general scheme (GS) model applied to $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$ with directions $\boldsymbol{g}_o = (\boldsymbol{g}_o^x, \boldsymbol{g}_o^y) \gneqq 0$ that may depend on $(\boldsymbol{x}_o, \boldsymbol{y}_o)$, we understand

$$(\text{GS})_o \qquad \min \theta \tag{11a}$$

$$\boldsymbol{X}\boldsymbol{\lambda} \leq \boldsymbol{x}_o + (\psi^x(\theta) - 1)\boldsymbol{g}_o^x, \tag{11b}$$

$$\boldsymbol{Y}\boldsymbol{\lambda} \geq \boldsymbol{y}_o + (\psi^y(\theta) - 1)\boldsymbol{g}_o^y, \tag{11c}$$

$$\boldsymbol{e}^T\boldsymbol{\lambda} = 1, \quad \boldsymbol{\lambda} \geq \boldsymbol{0}. \tag{11d}$$

Here, the real functions ψ, their domains (dom), and their images (im) satisfy the following assumptions.

(A1) $\text{dom}(\psi^x) = (a^x, \infty)$ with $a^x \in \{-\infty, 0\}$ and $\text{dom}(\psi^y) = (a^y, \infty)$ with $a^y \in \{-\infty, 0\}$;

(A2) ψ^x is smooth, concave, increasing, and ψ^y is smooth, convex, decreasing;

(A3) $\psi^x(1) = \psi^y(1) = 1$;

(A4) $\text{im}(\psi^x) = (b^x, \infty)$ with $b^x = -\infty$ if $\boldsymbol{g}^y = \boldsymbol{0}$ and $b^x \in \{-\infty, 0\}$ otherwise; $\text{im}(\psi^y) = (b^y, \infty)$ with $b^y \in \{-\infty, 0\}$.

Due to the monotonicity of ψ^x and ψ^y assumed in (A2), the assumption (A4) means $\lim_{\theta \to a_+^x} \psi^x(\theta) = b^x$ and $\lim_{\theta \to \infty} \psi^x(\theta) = \infty$ and $\lim_{\theta \to a_+^y} \psi^y(\theta) = \infty$ and $\lim_{\theta \to \infty} \psi^y(\theta) = b^y$.

If the vectors $\boldsymbol{g}_o^x$ and $\boldsymbol{g}_o^y$ are positive, then the GS model (11) operates in the full input-output space, and hence it is a graph model in the terminology of Färe et al. (1985). With $\boldsymbol{g}_o^x = 0$ or $\boldsymbol{g}_o^y = 0$,[6] one obtains output-oriented or input-oriented models, respectively.[7]

Denote the right-hand sides of (11b) and (11c) by $\boldsymbol{\phi}_o^x$ and $\boldsymbol{\phi}_o^y$, respectively, that is,

$$\boldsymbol{\phi}_o^x(\theta) := \boldsymbol{x}_o + (\psi^x(\theta) - 1)\boldsymbol{g}_o^x; \qquad \boldsymbol{\phi}_o^y(\theta) := \boldsymbol{y}_o + (\psi^y(\theta) - 1)\boldsymbol{g}_o^y. \tag{12}$$

---

[6]In the case $\boldsymbol{g}_o^x = 0$ or $\boldsymbol{g}_o^y = 0$ the contribution of ψ^x in (11b) or ψ^y in (11c) vanishes; in such cases we take $\text{dom}(\psi^x) = (-\infty, \infty)$ or $\text{dom}(\psi^y) = (-\infty, \infty)$, respectively.

[7]In the case of input/output-oriented GS models (where $\boldsymbol{g}_o^y = \boldsymbol{0}$ or $\boldsymbol{g}_o^x = \boldsymbol{0}$), the GS model can be converted to a DDF-g input/output-oriented model with the same projection as the GS model. This observation follows from the fact that the functions $\psi^x(\cdot)$ and $\psi^y(\cdot)$ are assumed to be monotonic so that their inverse functions exist and the transformations $\psi^x(\theta) =: a \iff (\psi^x)^{-1}(a) = \theta$ or $\psi^y(\theta) =: b \iff (\psi^y)^{-1}(b) = \theta$ can be applied. In the case of graph models, the inverse transformation allows one to convert the GS model into a form where the input (resp., output) side is linear in θ.

Model	$\boldsymbol{g}_o^x$	$\boldsymbol{\phi}_o^x(\theta)$	$\boldsymbol{g}_o^y$	$\boldsymbol{\phi}_o^y(\theta)$	ψ^x	ψ^y
BCC-I	$\boldsymbol{x}_o$	$\theta\boldsymbol{x}_o$	$\boldsymbol{0}$	$\boldsymbol{y}_o$	θ	ψ^y [8]
BCC-O	$\boldsymbol{0}$	$\boldsymbol{x}_o$	$\boldsymbol{y}_o$	$\frac{1}{\theta}\boldsymbol{y}_o$	ψ^x [9]	$\frac{1}{\theta}$
DDF-g	$\boldsymbol{g}_o^x$	$\boldsymbol{x}_o + (\theta-1)\boldsymbol{g}_o^x$	$\boldsymbol{g}_o^y$	$\boldsymbol{y}_o + (1-\theta)\boldsymbol{g}_o^y$	θ	$2-\theta$
DDF	$\boldsymbol{x}_o$	$\theta\boldsymbol{x}_o$	$\boldsymbol{y}_o$	$(2-\theta)\boldsymbol{y}_o$	θ	$2-\theta$
HDF	$\boldsymbol{x}_o$	$\theta\boldsymbol{x}_o$	$\boldsymbol{y}_o$	$\frac{1}{\theta}\boldsymbol{y}_o$	θ	$\frac{1}{\theta}$
HDF-g	$\boldsymbol{g}_o^x$	$\boldsymbol{x}_o + (\theta-1)\boldsymbol{g}_o^x$	$\boldsymbol{g}_o^y$	$\boldsymbol{y}_o + (\frac{1}{\theta}-1)\boldsymbol{g}_o^y$	θ	$\frac{1}{\theta}$

Table 2: Parameterization of the standard path-based models.

It is easy to see that the models presented in the previous section are special cases of the GS model. These models, taken in conjunction with the directions in Table 1 and the usual assumptions on the positiveness of the data (that is, condition (2) for the BCC models and (5) for the hyperbolic model), will be called *standard path-based models*. The corresponding parameterisations in Table 2 indicate that only two choices of the function ψ^y are associated with the standard models: linear, $\psi^y(\theta) = 2 - \theta$, and hyperbolic, $\psi^y(\theta) = \frac{1}{\theta}$. Only one form of ψ^x appears: $\psi^x(\theta) = \theta$.

However, other choices of ψ that satisfy the assumptions (A1)–(A4) are possible, for example $\psi^x(\theta) = \theta^{-p}$, $-1 < p < 0$, or $\psi^x(\theta) = 1 + \ln\theta$, and $\psi^y(\theta) = \theta^{-p}$, $p > 0$, $\psi^y(\theta) = 1 - \ln\theta$ with domains $(0, \infty)$, or $\psi^y(\theta) = \mathrm{e}^{1-\theta}$ with domain $(-\infty, \infty)$. We will be able to fine-tune the model properties by making suitable choices of ψ^x and ψ^y and directions. Note that the choice of functions $\psi^x(\theta) = \theta^{1-p}, \psi^y(\theta) = \theta^{-p}, p \in [0, 1]$ and directions $(\boldsymbol{g}^x, \boldsymbol{g}^y) = (\boldsymbol{x}_o, \boldsymbol{y}_o) > \boldsymbol{0}$ leads to the so-called generalised distance function introduced by Chavas and Cox (1999).

The next subsection shows that the $(\mathrm{GS})_o$ model is well defined as a convex programme, that is, for any $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$ the optimal value θ^* of $(\mathrm{GS})_o$ is finite and the minimum is attained. To this end, we introduce a geometric interpretation of the model.

3.1. Geometric interpretation

The definition of the technology set $\mathcal{T}$ in (1) and the notation in (12) allow us to rewrite $(\mathrm{GS})_o$ (11) in the form

$$\min\{\theta : (\boldsymbol{\phi}_o^x(\theta), \boldsymbol{\phi}_o^y(\theta)) \in \mathcal{T}\}. \tag{13}$$

In this context, we can interpret the $(\mathrm{GS})_o$ model as a path-based model, where the map

$$\boldsymbol{\phi}_o : \theta \mapsto (\boldsymbol{\phi}_o^x(\theta), \boldsymbol{\phi}_o^y(\theta)) \tag{14}$$

defines a continuous path in the input-output space $\mathbb{R}^m \times \mathbb{R}^s$ parameterised by $\theta \in \mathcal{D}$, where

$$\mathcal{D} = \mathrm{dom}(\psi^x) \cap \mathrm{dom}(\psi^y). \tag{15}$$

---

[8]Since $\boldsymbol{g}_o^y = 0$, the term $\boldsymbol{g}_o^y \psi^y$ vanishes.
[9]Since $\boldsymbol{g}_o^x = 0$, the term $\boldsymbol{g}_o^x \psi^x$ vanishes.
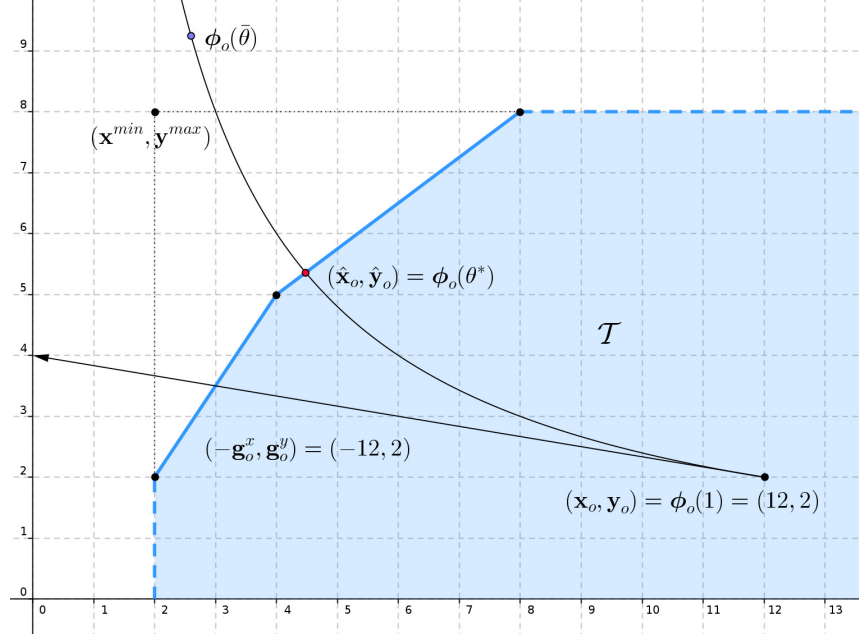
Figure 1: Hyperbolic path $\boldsymbol{\phi}_o(\theta)$ for direction $(12, 2)$ as an illustration of Theorem 1. The path is a smooth curve that starts for $\theta = 1$ at $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$ and for decreasing values of θ it moves in a set of points that gradually dominate each other, until the curve for some $\bar{\theta} < 1$ reaches the value $\phi_o^y(\bar{\theta}) > y^{\max}$. As a consequence, the path has to leave $\mathcal{T}$ at some $(\hat{\boldsymbol{x}}_o, \hat{\boldsymbol{y}}_o) \in \partial\mathcal{T}$ for some $\theta = \theta_o^*$.

It is seen from the definition of the path in (12) and (14) that each path is determined by a specific choice of $\boldsymbol{x}_o, \boldsymbol{y}_o, \boldsymbol{g}_o^x, \boldsymbol{g}_o^y$, and ψ^x, ψ^y (the *parameters of the path*). Unless otherwise stated, we take *standard parameter choices*, that is,

$$(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}, \quad (g_o^x, g_o^y) \gneq 0, \quad \text{and} \quad \psi^x, \psi^y \text{ that meet (A1)–(A4).} \tag{16}$$

The next theorem summarises the properties of the path $\boldsymbol{\phi}_o$ that are common for all standard parameter choices. Its proof is placed in Appendix A. The path properties are illustrated in Figure 1 with an example of a hyperbolic path.

Theorem 1. *The path $\boldsymbol{\phi}_o$ defined by (12) and (14) with standard parameter choices (16) has the following properties.*

(a) *For $\theta_1 \leq \theta_2 \in \mathcal{D}$, the point $\boldsymbol{\phi}_o(\theta_1)$ on the path $\boldsymbol{\phi}_o$ dominates the point $\boldsymbol{\phi}_o(\theta_2)$, (i.e., $\boldsymbol{\phi}_o(\theta_1) \succsim \boldsymbol{\phi}_o(\theta_2)$), and $\boldsymbol{\phi}_o(\theta_1) \neq \boldsymbol{\phi}_o(\theta_2)$, if and only if $\theta_1 < \theta_2$.*

(b) *$\boldsymbol{\phi}_o(1) = (\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$.*

(c) *There exists i such that $\lim_{\theta \to a_-^x} \phi_{io}^x(\theta) = -\infty$ or r such that $\lim_{\theta \to a_+^y} \phi_{ro}^y(\theta) = +\infty$ (see Figure 1 for illustration).*

(d) *There exists $\theta_o^* \leq 1$ such that $\boldsymbol{\phi}_o(\theta^*) \in \mathcal{T}$ and $\boldsymbol{\phi}_o(\theta) \notin \mathcal{T}$ for all $\theta < \theta^*$. As a consequence, the point $\boldsymbol{\phi}_o(\theta_o^*)$ belongs to the boundary of $\mathcal{T}$ and θ^* is the finite optimal value of the $(GS)_o$ model (11).*

13

By Theorem 1, the GS model applied to $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$ with directions $(\boldsymbol{g}_o^x, \boldsymbol{g}_o^y) \gneqq 0$ is well defined. The optimal value θ_o^* is called *efficiency score*, or alternatively *value of the efficiency measure* for $(\boldsymbol{x}_o, \boldsymbol{y}_o)$. The point $(\boldsymbol{\phi}_o^x(\theta_o^*), \boldsymbol{\phi}_o^y(\theta_o^*))$ on the path $\boldsymbol{\phi}_o$ is called the *projection* of $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ in the $(\text{GS})_o$ model. Hereafter, we shall denote the projection point more compactly by

$$(\hat{\boldsymbol{x}}_o, \hat{\boldsymbol{y}}_o) := (\boldsymbol{\phi}_o^x(\theta_o^*), \boldsymbol{\phi}_o^y(\theta_o^*)). \tag{17}$$

Note that although the projection point $(\hat{\boldsymbol{x}}_o, \hat{\boldsymbol{y}}_o) \in \partial\mathcal{T}$ is defined uniquely, the uniqueness may not extend to the $\boldsymbol{\lambda}$-component of an optimal solution $(\boldsymbol{\lambda}^*, \theta^*)$. Consequently, the program $(\text{GS})_o$ may have multiple optimal slacks given by

$$(\boldsymbol{s}^{x*}, \boldsymbol{s}^{y*}) := (\hat{\boldsymbol{x}}_o - \boldsymbol{X}\boldsymbol{\lambda}^*, \boldsymbol{Y}\boldsymbol{\lambda}^* - \hat{\boldsymbol{y}}_o) \geq \boldsymbol{0}.$$

Lemma 2. *Let $(\boldsymbol{\lambda}_o^*, \theta_o^*)$ be an optimal solution of $(GS)_o$. Then at least one of the inequalities in (11b) or (11c) is tight, i.e, at least one component of the corresponding slacks is zero.*

Proof. If all components of the slacks are positive, then due to the assumed continuity of the path there is $\epsilon > 0$ such that $(\boldsymbol{\lambda}_o^*, \theta_o^* - \epsilon)$ is feasible, which contradicts the optimality of θ_o^*. $\qquad\square$

4. Properties of the general model

We will now analyse the $(\text{GS})_o$ model in the light of the ten desirable properties (P1)–(P9). The only property that the general scheme $(\text{GS})_o$ satisfies regardless of the specific choice of directions is the unique projection property (P1). This is a simple consequence of Theorem 1. The remaining properties depend on the choice of the direction vector g_o and the functions ψ^x and ψ^y. We examine them one at a time in the following subsections.

4.1. Indication

(P2) (a) If the evaluated unit is strongly efficient, then the measure equals 1.

(b) If the measure equals 1, then the evaluated unit is strongly efficient.

The part (a) of this property holds universally in the (GS) scheme as indicated in the following theorem.

Theorem 2. *If $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ is strongly efficient in $\mathcal{T}$, then $\theta_o^* = 1$.*

Proof. If $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ is strongly efficient, then the set of points from $\mathcal{T}$ dominating $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ contains only $(\boldsymbol{x}_o, \boldsymbol{y}_o)$. In view of $\boldsymbol{\phi}_o(1) = (\boldsymbol{x}_o, \boldsymbol{y}_o)$ and Theorem 1(a), one has $\theta_o^* = 1$. $\qquad\square$

Note that $\theta_o^* = 1$ in conjunction with Theorem 1(d) only yield $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \partial\mathcal{T}$, hence one must admit the possibility that $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ belongs to $\partial^W\mathcal{T}$ in some cases with $\theta_o^* = 1$.

Theorem 3. *The GS model with positive directions violates the part (b) of (P2) for any $\mathcal{T}$.*

Proof. Let $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \partial^W \mathcal{T}$ and $\boldsymbol{g}_o > \boldsymbol{0}$. We want to show that $\theta_o^* = 1$. Assume by contradiction that $\theta_o^* < 1$. Then $\hat{\boldsymbol{x}}_o = \boldsymbol{x}_o + (\psi^x(\theta_o^*) - 1)\boldsymbol{g}_o^x < \boldsymbol{x}_o$ and $\hat{\boldsymbol{y}}_o = \boldsymbol{y}_o + (\psi^y(\theta_o^*) - 1)\boldsymbol{g}_o^y > \boldsymbol{y}_o$, and $(\hat{\boldsymbol{x}}_o, \hat{\boldsymbol{y}}_o) \in \mathcal{T}$, which is, by Lemma 1(d), in contradiction with the weak but not strong efficiency of $(\boldsymbol{x}_o, \boldsymbol{y}_o)$. □
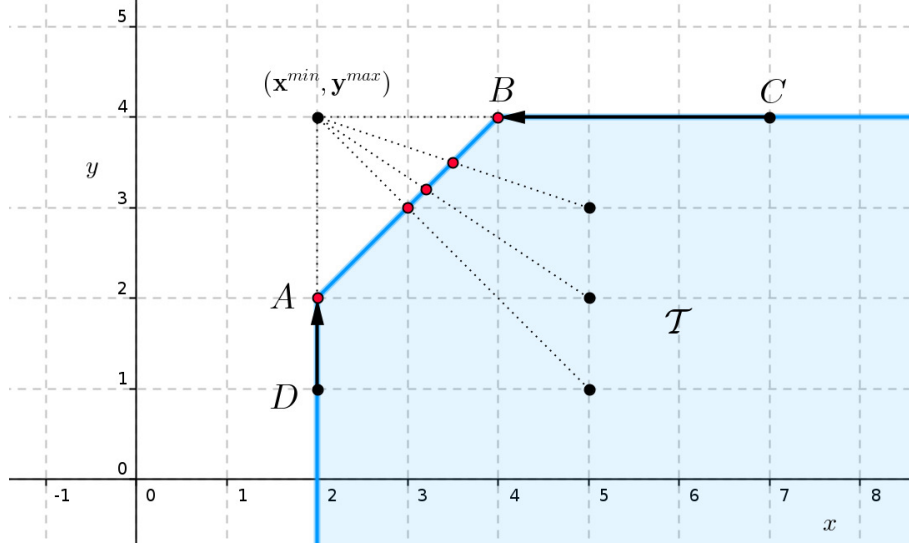

Figure 2: In the two dimensional technology set $\mathcal{T}$, the DDF-(G2) model projects units from $\mathcal{T}$ onto $\partial^S \mathcal{T}$. The set $\mathcal{T}$ is generated by four units. The units A and B are strongly efficient, while C and D belong to $\partial^W \mathcal{T}$. The (G2) direction for C is $(5, 0)$ and it points from C to B which implies that $\theta_C^* < 1$. The (G2) direction for D is $(0, 1)$ and it points from D to A which implies that $\theta_D^* < 1$. Similar observation applies to all units from $\partial^W \mathcal{T}$.

Remark 2. By Theorem 3, all standard path-based models with directions (G1) and (G3)–(G6) violate the part (b) of (P2) for any $\mathcal{T}$. In the case of DDF-(G2), whether (P2) (b) holds depends on the data configuration (the shape of $\mathcal{T}$). It holds in Figure 2, where each unit $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \partial^W \mathcal{T}$ is projected onto $\partial^S \mathcal{T}$ in DDF-(G2) model, and hence $\theta_o^* < 1$. A failure of (P2) (b) is seen in Figure 3, where the unit C belongs to $\partial^W \mathcal{T}$ and the corresponding paths at C point out of $\mathcal{T}$. This implies that the efficiency scores of C is equal to one. On the other hand, HDF-(G2) violates part (b) of (P2) even in the case of technology set depicted in Figure 2. Indeed, since $\frac{1}{\theta} - 1 > 1 - \theta$ for $\theta < 1$, the hyperbolic path for C is above the directional path for C for $\theta < 1$, and hence leaves $\mathcal{T}$ at C.

The tools to determine whether a unit with an efficiency score equal to one is strongly efficient will be presented in the next subsection, where this problem will be solved in a more general setting.

4.2. Strong efficiency of projection

(P3) The projection point generated by the measure is strongly efficient.

The projection in the $(GS)_o$ model is located on the boundary of $\mathcal{T}$ but not necessarily on its strongly efficient part. Note that the lack of strong efficiency of a projection is a known possibility
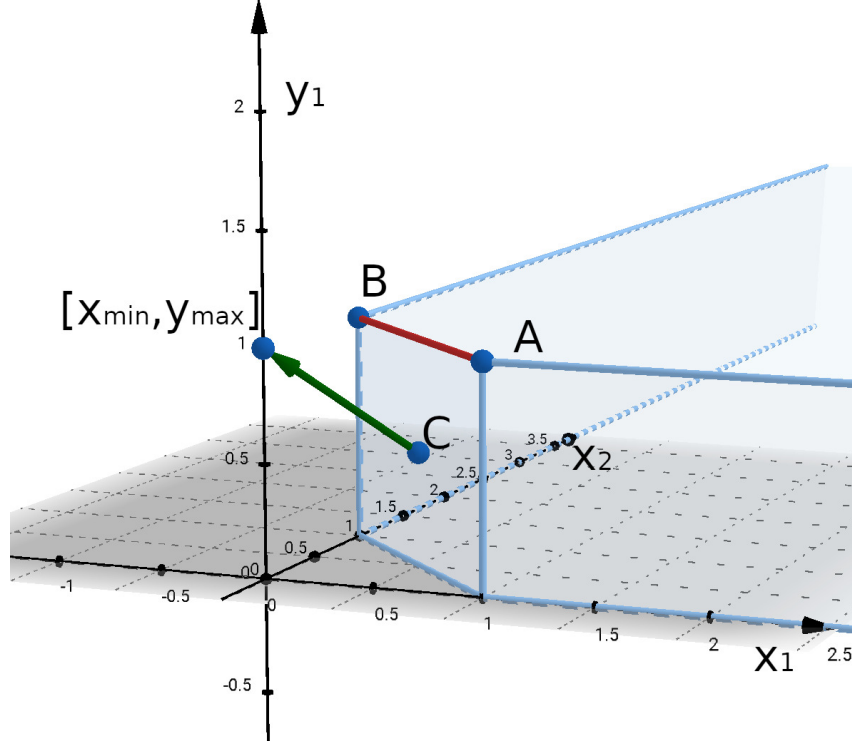
Figure 3: An example of $\mathcal{T}$, for which the DDF-(G2) model does not project $C \in \partial^W \mathcal{T}$ onto the strongly efficient frontier $\partial^S \mathcal{T} = |AB|$. The two-input and single-output technology set is generated by two strongly efficient units, $A = (1, 0, 1)$ and $B = (0, 1, 1)$. The unit $C = (0.5, 0.5, 0.5)$ is weakly but not strongly efficient. Here $(\boldsymbol{x}^{\min}, \boldsymbol{y}^{\max}) = (0, 0, 1)$ and (G2) direction for C is $\boldsymbol{g}_C = (0.5, 0.5, 0.5)$. Since $\boldsymbol{g}_C > 0$, Theorem 3 yields that the part (b) of (P2) is not satisfied. The DDF-(G2) path for C with the $\boldsymbol{g}_C$ direction leaves $\mathcal{T}$ at C, which means that C is projected onto itself with $\theta_C^* = 1$.

in the BCC and DDF models. In this subsection, we show that this is the case for all standard path-based models.

The absence of strong efficiency of the projection has knock-on effects, such as the lack of strict monotonicity of the efficiency score or the failure of one-to-one identification of efficiency. Moreover, if a unit is not projected onto the strongly efficient frontier, then its efficiency score does not capture all sources of inefficiencies in some input/output components, and hence it overestimates the unit's performance. Therefore, it is important to recognise whether the specific unit was projected onto a strongly efficient frontier or not. Unfortunately, this recognition cannot be made from the mere value of the optimal score. However, in connection with BCC and DDF, tools have been developed to detect this feature. Now we are adopting them into our scheme.

The next theorem states that to identify the strong efficiency of projections, one needs to know more than just the optimal value. The proof is located in Appendix A.

Theorem 4. *The projection of $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ is strongly efficient if and only if for each optimal solution $(\boldsymbol{\lambda}_o^*, \theta_o^*)$ of (11), the inequality constraints (11b) and (11c) are satisfied with equality.*

16

Theorem 4 identifies whether the projection is from the strongly efficient frontier theoretically only. A practical method to identify the strong efficiency of a weakly efficient unit $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$ with $\theta_o^* = 1$, but also the projection of an inefficient unit $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$ with $\theta_o^* < 1$, is called *second phase procedure*. The procedure solves a modified programme that, at a fixed value of θ_o^*, maximises the sum of slacks. In the context of the GS scheme, it reads as follows:

$$
\begin{aligned}
\max \quad & \boldsymbol{e}^T \boldsymbol{s}^x + \boldsymbol{e}^T \boldsymbol{s}^y \\
& \boldsymbol{X}\boldsymbol{\lambda} + \boldsymbol{s}^x = \boldsymbol{\phi}_o^x(\theta_o^*), \\
& \boldsymbol{Y}\boldsymbol{\lambda} - \boldsymbol{s}^y = \boldsymbol{\phi}_o^y(\theta_o^*), \\
& \boldsymbol{e}^T\boldsymbol{\lambda} = 1, \ \boldsymbol{\lambda} \geq \boldsymbol{0}, \boldsymbol{s}^x \geq \boldsymbol{0}, \boldsymbol{s}^y \geq \boldsymbol{0}.
\end{aligned}
\tag{18}
$$

Theorem 5. *The optimal value in* (18) *vanishes if and only if the projection of* $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ *in* $(GS)_o$ *is strongly efficient. Furthermore, for each optimal solution* $(\boldsymbol{\lambda}^*, \boldsymbol{s}^{x*}, \boldsymbol{s}^{y*})$ *in* (18), *the unit* $(\boldsymbol{X}\boldsymbol{\lambda}^*, \boldsymbol{Y}\boldsymbol{\lambda}^*)$ *is a strongly efficient benchmark for* $(\boldsymbol{x}_o, \boldsymbol{y}_o)$.

Proof. The first part follows from Theorem 4. The second part claims that $(\boldsymbol{X}\boldsymbol{\lambda}^*, \boldsymbol{Y}\boldsymbol{\lambda}^*) = (\hat{\boldsymbol{x}}_o - \boldsymbol{s}^{x*}, \hat{\boldsymbol{y}}_o + \boldsymbol{s}^{y*}) \in \partial^S \mathcal{T}$. Suppose this is not the case. Then by Lemma 1(c), there exists a unit in $\mathcal{T}$ (corresponding to some $\hat{\boldsymbol{\lambda}} \geq \boldsymbol{0}$ with $\boldsymbol{e}^T\hat{\boldsymbol{\lambda}} = 1$) that dominates but is not equal to $(\boldsymbol{X}\boldsymbol{\lambda}^*, \boldsymbol{Y}\boldsymbol{\lambda}^*)$. The slacks of $(\hat{\boldsymbol{x}}_o, \hat{\boldsymbol{y}}_o)$ relative to the new unit have increased, which contradicts the optimality of the original slacks in (18). $\square$

Remark 3. While in the class of slacks-based models the notions of 'projection' and 'benchmark' coincide (see Remark 2 in Halická and Trnovská, 2021), in the class of path-based models it is necessary to single out one projection from among multiple benchmarks. The projection is uniquely determined by (17) as the point (benchmark) on the path, where the path leaves $\partial \mathcal{T}$, which also pins down the optimal value of the objective function. However, the model is capable of generating other benchmarks, i.e., points $(\boldsymbol{X}\boldsymbol{\lambda}^*, \boldsymbol{Y}\boldsymbol{\lambda}^*) \in \partial T$ that correspond to the alternate optimal solutions $(\boldsymbol{\lambda}^*, \theta_o^*)$ of $(GS)_o$. Some of them lie in $\partial^S \mathcal{T}$ and can be generated by optimal solutions of (18).

Remark 4. Property (P3) implies (P2). Indeed, if $\theta_o^* = 1$, then $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ coincides with its projection. Now, if (P3) holds, this projection is strongly efficient, which yields the (b) part of (P2). The part (a) of (P2) holds universally by Theorem 2.

As a consequence of the implication (P3) $\Rightarrow$ (P2), the conclusions of Remark 2 regarding failure of (P2) for standard path-based models carry over to property (P3).

4.3. Boundedness

(P4) The measure takes values between zero and one.

If $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$, then automatically $\theta_o^* \leq 1$ (since $\theta = 1$ is a feasible solution). On the other hand, the property $\theta_o^* \geq 0$ is based on the domains of ψ^x and ψ^y. With $\text{dom}(\psi^x) = (0, \infty)$ or $\text{dom}(\psi^y) = (0, \infty)$, which only occur for non-linear ψ^x and ψ^y, one has $\theta_o^* > 0$ for any choice of g_o. With linear ψ one has $\text{dom}(\psi) = \mathbb{R}$, and hence (P4) may fail. In the following theorem, we formulate sufficient conditions on g_o for (P4) to hold.

Theorem 6. *Consider a GS model with $\psi^x(\theta) = \theta$, $\psi^y(\theta) = 2 - \theta$, and directions $g_o \gneq 0$. If*

$$\max_i \left\{ \frac{x_i^{\min} + g_{io}^x - x_{io}}{g_{io}^x} : g_{io}^x > 0 \right\} \geq 0 \quad or \quad \max_r \left\{ \frac{y_{ro} + g_{ro}^y - y_r^{\max}}{g_{ro}^y} : g_{ro}^y > 0 \right\} \geq 0,$$

then $\theta_o^* \geq 0$.

Proof. We proceed by constructing lower bounds for the optimal score θ_o^*. If $g_{io}^x > 0$, then the i-th component of (11b) yields $\frac{x_i^{\min} + g_{io}^x - x_{io}}{g_{io}^x} \leq \psi^x(\theta_o^*) = \theta_o^*$. Similarly, if $g_{ro}^y > 0$, then the r-th component of (11c) yields $\frac{y_r^{\max} + g_{ro}^y - y_{ro}}{g_{ro}^y} \geq \psi^y(\theta_o^*) = 2 - \theta_o^*$ and hence $\frac{y_{ro} + g_{ro}^y - y_r^{\max}}{g_{ro}^y} \leq \theta_o^*$. If the maximal lower bound is non-negative, then $\theta_o^* \geq 0$, which completes the proof. □

Sufficient conditions for other linear ψ^x and ψ^y follow by an obvious modification of the proof.

Remark 5. Theorem 6 offers an easy route to verify that DDF-g with directions (G1)–(G3) satisfies boundedness (P4). A counterexample with negative θ_o^* for directions (G4)–(G6) is provided by a single-input, single-output model with two units $A = (1, 5)$ and $B = (5, 1)$. Here, the average direction (G4) is $(3, 3)$ and the standard deviation direction (G5) equals $(2, 2)$. The scores for the non-efficient unit B are $-1/3$, -1, and -3, respectively, for the directions (G4)–(G6). The HDF-g family (including HDF itself) satisfies (P4) trivially thanks to the non-negative domain of ψ.

The results of Theorem 6 are in line with the partial results of Sahoo et al. (2014) related to the input-oriented directional distance model.

4.4. Unit invariance

(P5) The value of the measure does not depend on the units of measurement in the input and output variables.

The next theorem provides a necessary and sufficient condition for unit invariance in terms of invariance of directions with respect to the data transformation. Its proof is given in Appendix A.

Theorem 7. *Let $\boldsymbol{C} \in \mathbb{D}_{++}^m$ and $\boldsymbol{B} \in \mathbb{D}_{++}^s$ be diagonal matrices with positive elements on the diagonal. For all $j = 1, \ldots, n$, consider transformed inputs $\boldsymbol{C}\boldsymbol{x}_j$ and transformed outputs $\boldsymbol{B}\boldsymbol{y}_j$, respectively. Let $\boldsymbol{g}_o^x$, $\boldsymbol{g}_o^y$, and $\boldsymbol{g}_o^{Cx}$, $\boldsymbol{g}_o^{By}$ denote the directional vectors before and after the data transformation, respectively. The $(GS)_o$ model is unit invariant if and only if one has $\boldsymbol{g}_o^{Cx} = \boldsymbol{C}\boldsymbol{g}_o^x$ and $\boldsymbol{g}_o^{By} = \boldsymbol{B}\boldsymbol{g}_o^y$.*

Remark 6. Theorem 7 yields easily that GS models with (G1)–(G5) directions satisfy unit invariance (P5) but this property fails with directions (G6).

4.5. Translation invariance

(P6) The value of the measure is not affected by translation in inputs or outputs.

The next theorem provides a necessary and sufficient condition for translation invariance in terms of translation invariance of directions. Its proof is given in Appendix A.

Theorem 8. *The $(GS)_o$ model is translation invariant if and only if the directional vectors $\boldsymbol{g}_o^x$ and $\boldsymbol{g}_o^y$ are invariant to the translation of inputs and outputs, respectively.*

The translation invariance of the DDF-g models is analysed in detail by Aparicio et al. (2016). This subsection is consistent with their results and extends them to the whole class of GS models.

Remark 7. Using Theorem 8, one easily verifies that GS models with directions (G2), (G3), (G5), and (G6) satisfy translation invariance (P6) but those with directions (G1) and (G4) do not.

4.6. Monotonicity

(P7) An increase in any input or a decrease in any output relative to the evaluated unit, keeping other inputs and outputs constant, reduces or maintains the value of the measure.

For any fixed $\bar{\theta} \in \mathcal{D}$, we consider a function $\boldsymbol{\phi}(\bar{\theta})$: $(\boldsymbol{x}_o, \boldsymbol{y}_o) \mapsto \boldsymbol{\phi}_o(\bar{\theta})$ mapping each point $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$ to the point $(\boldsymbol{\phi}_o^x(\bar{\theta}), (\boldsymbol{\phi}_o^y(\bar{\theta}))$ on the path $\boldsymbol{\phi}_o(\theta)$ at $\theta = \bar{\theta}$. The monotonicity property of the models states that for any $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ and $(\boldsymbol{x}_q, \boldsymbol{y}_q)$ in $\mathcal{T}$ one has

$$(\boldsymbol{x}_o, \boldsymbol{y}_o) \succsim (\boldsymbol{x}_q, \boldsymbol{y}_q) \;\Rightarrow\; \theta_o^* \geq \theta_q^*.$$

The monotonicity of a model is linked with the monotonicity of $\boldsymbol{\phi}(\theta)$, which we define as follows.

Definition 1. *We say that $\boldsymbol{\phi}(\bar{\theta})$ is monotone on $\mathcal{T}$ in $\bar{\theta} \in \mathcal{D}$ if for any two units $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ and $(\boldsymbol{x}_q, \boldsymbol{y}_q)$ in $\mathcal{T}$, one has*

$$(\boldsymbol{x}_o, \boldsymbol{y}_o) \succsim (\boldsymbol{x}_q, \boldsymbol{y}_q) \;\Rightarrow\; \boldsymbol{\phi}_o(\bar{\theta}) \succsim \boldsymbol{\phi}_q(\bar{\theta}). \tag{19}$$

For a function $\boldsymbol{\phi}(\bar{\theta})$ such that $[\boldsymbol{\phi}^x(\bar{\theta})]_i$ depends only on $\boldsymbol{x}_{io}$ and $[\boldsymbol{\phi}^y(\bar{\theta})]_r$ depends only on $\boldsymbol{y}_{ro}$, the monotonicity property in Definition 19 simply means that $[\boldsymbol{\phi}^x(\bar{\theta})]_i$ and $[\boldsymbol{\phi}^y(\bar{\theta})]_r$ are nondecreasing in $\boldsymbol{x}_{io}$ and $\boldsymbol{y}_{ro}$, respectively.

Lemma 3. *Let $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ and $(\boldsymbol{x}_q, \boldsymbol{y}_q)$ be two units in $\mathcal{T}$ with the corresponding optimal values θ_o^*, θ_q^*. If $\boldsymbol{\phi}_o(\theta_o^*) \succsim \boldsymbol{\phi}_q(\theta_o^*)$, then each optimal solution $(\theta_o^*, \boldsymbol{\lambda}_o^*)$ of $(GS)_o$ is a feasible solution of $(GS)_q$. Furthermore, $\theta_q^* \leq \theta_o^*$.*

Proof. The assumptions of the lemma and feasibility of $(\theta_o^*, \boldsymbol{\lambda}_o^*)$ for $(GS)_o$ yield

$$\boldsymbol{X}\boldsymbol{\lambda}_o^* \leq \boldsymbol{\phi}_o^x(\theta_o^*) \leq \boldsymbol{\phi}_q^x(\theta_o^*), \quad \boldsymbol{Y}\boldsymbol{\lambda}_o^* \geq \boldsymbol{\phi}_o^y(\theta_o^*) \geq \boldsymbol{\phi}_q^y(\theta_o^*).$$

Therefore $(\theta_o^*, \boldsymbol{\lambda}_o^*)$ is also feasible for $(GS)_q$. Since θ_q^* is the optimal value for the minimisation problem $(GS)_q$, we get $\theta_q^* \leq \theta_o^*$. □

The next theorem is a consequence of Lemma 3. Here, we assume that the assumption of monotonicity is satisfied only at $\bar{\theta} \in \tilde{\mathcal{D}}$, where $\tilde{\mathcal{D}}$ is a set of those $\bar{\theta}$ for which there exists $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$ such that $\theta_o^* = \bar{\theta}$.

Theorem 9. *Suppose that $\boldsymbol{\phi}(\bar{\theta})$ is monotone at any $\bar{\theta} \in \tilde{\mathcal{D}}$. Then the GS model satisfies the property of monotonicity (P7).*

Remark 8. Note that if $\boldsymbol{g}_o$ does not depend on $(\boldsymbol{x}_o, \boldsymbol{y}_o)$, then $\boldsymbol{\phi}(\theta)$ is monotone for any choice of ψ^x and ψ^y in any $\theta \in \mathcal{D}$. Therefore, by Theorem 9, (GS) models with directions (G3)–(G6) satisfy the property (P7). For the (G1) direction with arbitrary (standard) ψ^x and ψ^y, the monotonicity assumption of Theorem 9 is satisfied in $\bar{\theta} \in \mathcal{D}$, provided that $\psi^x(\bar{\theta}) > 0$ and $\psi^y(\bar{\theta}) > 0$ hold. These conditions are satisfied by the DDF-(G1) and HDF-(G1) models, since, by Remark 5, in these models θ^* reaches only positive values on $\mathcal{T}$. Therefore, both models meet the property (P7). For the (G2) direction with arbitrary (standard) ψ^x and ψ^y, the monotonicity holds at $\bar{\theta} \in \mathcal{D}$ provided that $\psi^x(\bar{\theta}) > 0$ and $\psi^y(\bar{\theta}) < 2$ hold. These conditions are again satisfied by both DDF-(G2) and HDF-(G2). In fact, according to Remark 5, DDF-(G2) meets (P4), and thus $\theta_o^* > 0$. Furthermore, $\theta_o^* \geq \frac{1}{2}$ on $\mathcal{T}$ in HDF-(G2) with easy computation.

In conclusion, all standard path-based models have the property of monotonicity (P7).

4.7. Strict monotonicity

(P8) An increase in any input or a decrease in any output relative to the evaluated unit, keeping other inputs as well as outputs constant, reduces the value of the measure.

We begin with a necessary condition for strict monotonicity (P8).

Theorem 10. *If there exists $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \partial^W \mathcal{T}$ such that $\theta_o^* = 1$, then the GS model does not meet the property of strict monotonicity.*

Proof. Since $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ is not a strongly efficient unit, by Theorem 5 there exists an optimal solution $(\boldsymbol{\lambda}^*, \boldsymbol{s}^{x*}, \boldsymbol{s}^{y*})$, of (18) such that $(\boldsymbol{s}^{x*}, \boldsymbol{s}^{y*}) \neq \boldsymbol{0}$ and $(\boldsymbol{X}\boldsymbol{\lambda}^*, \boldsymbol{Y}\boldsymbol{\lambda}^*) \in \partial^S \mathcal{T}$. Since $(\boldsymbol{x}_p, \boldsymbol{y}_p) = (\boldsymbol{X}\boldsymbol{\lambda}^*, \boldsymbol{Y}\boldsymbol{\lambda}^*)$ is strongly efficient, Theorem 2 yields $\theta_p^* = 1$. Strict monotonicity is violated because $(\boldsymbol{x}_p, \boldsymbol{y}_p)$ dominates and is different from $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ but $\theta_p^* = \theta_o^* = 1$. $\qquad\square$

We have shown in Theorem 3 that positive directions rule out the validity of (P2). As a consequence of Theorem 10, an analogous statement applies to (P8).

Theorem 11. *The GS model with positive directions fails in* (P8) *for any $\mathcal{T}$ defined by* (1).

Remark 9. According to Remark 2, all standard path-based models admit $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \partial^W \mathcal{T}$ such that $\theta_o^* = 1$ for some data configurations and hence by Theorem 10 do not satisfy (P8).

4.8. Super-efficiency

(P9) The value of the measure of a unit outside the technology set is well defined and finite.

It is well understood that for oriented models super-efficiency may fail for *some* $(\boldsymbol{x}_o, \boldsymbol{y}_o)$. Specifically, if $(\boldsymbol{x}_o, \boldsymbol{y}_o) \notin \mathcal{T}$, some of the inequalities in (11b) or (11c) are not satisfied for $\theta = 1$. To remedy the situation, one must increase the value of $\theta = 1$ and thus increase $\phi_o^x(\theta)$ on the input side and decrease $\phi_o^y(\theta)$ on the output side. This increase/decrease is impossible if any of the components of $\phi_o^x(\theta)$ or $\phi_o^y(\theta)$ are not dependent on θ. It follows that if one wishes to ensure (P9) for arbitrary $(\boldsymbol{x}_o, \boldsymbol{y}_o) \notin \mathcal{T}$, one must consider only the graph models where $\boldsymbol{g}_o^x > 0$ and $\boldsymbol{g}_o^y > 0$. The following

theorem establishes some necessary and some sufficient conditions for (P9). Its proof is located in Appendix A.

Theorem 12. *For $(\boldsymbol{x}_o, \boldsymbol{y}_o) \notin \mathcal{T}$ and $(\boldsymbol{g}_o^x, \boldsymbol{g}_o^y) \gneqq 0$, the following statements hold.*

(a) *If $(GS)_o$ is feasible, that is, if there exists $\bar{\theta} \in \mathrm{dom}(\psi)$ such that $(\boldsymbol{\phi}_o^x(\bar{\theta}), \boldsymbol{\phi}_o^y(\bar{\theta})) \in \mathcal{T}$, then $(GS)_o$ admits an optimal solution and the optimal value satisfies $\theta_o^* > 1$.*

(b) *Let $\boldsymbol{g}_o > 0$. If $\boldsymbol{y}^{\min} - \boldsymbol{y}_o + \boldsymbol{g}_o^y > 0$ or $\mathrm{im}(\psi^y) = (-\infty, \infty)$, then $(GS)_o$ admits an optimal solution and the optimal value satisfies $\theta_o^* > 1$.*

(c) *Assume $\mathrm{im}(\psi^y) = (0, \infty)$ and let $\boldsymbol{g}_o > 0$. If (i) there is no $\boldsymbol{\lambda} \geq 0$ with $\boldsymbol{e}^T \boldsymbol{\lambda} = 1$ such that inequality $\boldsymbol{Y}\boldsymbol{\lambda} \geq \boldsymbol{y}_o - \boldsymbol{g}_o^y$ is satisfied or (ii) $y_r^{\max} - y_{ro} + g_{ro}^y < 0$ for some r, then $(GS)_o$ is infeasible.*

Remark 10. With $\boldsymbol{g}_o > 0$, DDF-g models are feasible for all units outside of the technology set by Theorem 12(b) ($\mathrm{dom}(\psi) = (-\infty, \infty)$). As a result, the entire family of DDF-g models satisfies the property (P9) provided that the directional vectors are positive. Note that the positivity of the directions (G2) for $(\boldsymbol{x}_o, \boldsymbol{y}_o) \notin \mathcal{T}$ is satisfied only if $\boldsymbol{x}_o > \boldsymbol{x}^{\min}$ and $\boldsymbol{y}_o < \boldsymbol{y}^{\max}$.

In the class of HDF-g models over positive data, only the HDF model with positive (G1) directions is super-efficient since it satisfies the assumptions of Theorem 12(b).

The HDF-g models over positive data with (G3)–(G6) directions do not guarantee (P9). In fact, one can always choose $(\boldsymbol{x}_o, \boldsymbol{y}_o) \notin \mathcal{T}$ with $y_{ro} > y_r^{\max} + g_{ro}$ for some r. The infeasibility of the corresponding $(GS)_o$ then follows from (ii) of Theorem 12(c).

Condition (ii) is too coarse for HDF-(G2) because inequality $y_{ro} > y_r^{\max} + g_{ro}$ is incompatible with positive (G2) directions. Instead, we shall provide an example that meets condition (i) of Theorem 12(c). Consider a one-input, two-output production technology with the technology set generated by two units $A = (1, 1, 10)$ and $B = (1, 10, 1)$. It can be easily shown that the unit $(\boldsymbol{x}_o, \boldsymbol{y}_o) = (2, 8, 8)$ is outside $\mathcal{T}$. The corresponding (G2) direction is equal to $\boldsymbol{g}_o = (1, 2, 2)$. An easy calculation now shows that condition (i) of Theorem 12(c) is satisfied, and hence HDF-g with (G2) directions does not guarantee super-efficiency (P9).

Remark 11. The definition of the technology set $\mathcal{T}$ in (1), even when applied to non-negative data, does not presuppose non-negative values of outputs. Consequently, for the purposes of super-efficiency measurement, we allow projections of units $0 \leq (\boldsymbol{x}_o, \boldsymbol{y}_o) \notin \mathcal{T}$ onto elements of the frontier with negative outputs. Therefore, our results differ from those of studies by other authors, where projections onto the negative part of the frontier are treated by definition as *infeasible*.

For example, in Briec and Kerstens (2009) (in the context of general economic productivity theories), DDF-g is classified as not having property (P9) only because some units project onto the negative part of the frontier. To rule out such a source of infeasibility, the authors separately consider models, where the technological set is extended by the so-called *free disposal cone*, which then matches our approach. Johnson and McGinnis (2009) praise the HDF model, which in contrast to the DDF-g models is always feasible for positive data. This can also be seen in our GS scheme,

since the HDF projection with (G1) directions reads $(\theta\boldsymbol{x}_o, \frac{1}{\theta}\boldsymbol{y}_o)$ and thus is positive for positive data. Mehdiloozad and Roshdi (2014) provide a comprehensive review of super-efficient DEA models and analyse the pivotal role of the directional vector for super-efficiency.

4.9. Homogeneity

(P10) **Homogeneity.** Feasible scaling of the input vector and the output vector of the evaluated unit by a power of $\mu > 0$ (the power may be different for inputs and outputs) results in the efficiency score being scaled by a power of μ.

In this section, the efficiency score provided by the GS model for a point $(x, y) \in \mathcal{T}$ will be denoted by $\theta^*(\boldsymbol{x}, \boldsymbol{y})$. Here, the argument in θ^* indicates that the efficiency score depends on $(\boldsymbol{x}, \boldsymbol{y})$ not only directly, but also indirectly through unit-dependent directional vectors $\boldsymbol{g}$.

The next definition is consistent with the definition of the so-called almost homogeneity used in connection with the measure of hyperbolic efficiency in Cuesta and Zofío (2005).

Definition 2. *We say that the measure θ^* generated by a GS model over a technology set $\mathcal{T}$ is homogeneous of degree (α, β, γ) if the relation*

$$\theta^*(\mu^\alpha \boldsymbol{x}, \mu^\beta \boldsymbol{y}) = \mu^\gamma \theta^*(\boldsymbol{x}, \boldsymbol{y}) \tag{20}$$

holds for each $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{T}$ and for each $\mu > 0$ such that the value $\theta^*(\mu^\alpha \boldsymbol{x}, \mu^\beta \boldsymbol{y})$ is well defined.

Note that a measure is homogeneous of degree (α, β, γ), if and only if it is also homogeneous of degree $(k\alpha, k\beta, k\gamma)$ for any $k \neq 0$. To see this, it suffices to set $\bar{\mu} = \mu^k$ in (20) (the map $\mu \to \mu^k$ is one-to-one for $\mu > 0$). Therefore, without loss of generality, we consider only the homogeneity of degree $(\alpha, \beta, 1)$.

Remark 12. By Definition 2, $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{T}$, and therefore $\theta^*(\boldsymbol{x}, \boldsymbol{y})$ is well defined (Theorem 1(d)), that is, the minimum in (11) is attained as a finite value. However, the scaled point $(\mu^\alpha \boldsymbol{x}, \mu^\beta \boldsymbol{y})$ is not necessarily in $\mathcal{T}$. If the programme (11) for $(\mu^\alpha \boldsymbol{x}, \mu^\beta \boldsymbol{y})$ is not feasible, then $\theta^*(\mu^\alpha \boldsymbol{x}, \mu^\beta \boldsymbol{y}) = +\infty$, and (20) is not satisfied. On the other hand, if the programme (11) for $(\mu^\alpha \boldsymbol{x}, \mu^\beta \boldsymbol{y})$ is feasible, then $\theta^*(\mu^\alpha \boldsymbol{x}, \mu^\beta \boldsymbol{y})$ is well defined (Theorem 12(a)).

For the purposes of analysing the homogeneity property, we will restrict ourselves to the technology set $\mathcal{T}_+ = \mathcal{T} \cap (\mathbb{R}_+^m \times \mathbb{R}_+^s)$, generated with positive data. We will say that the GS model is homogeneous of degree $(\alpha, \beta, 1)$ if the corresponding measures θ^* are homogeneous of degree $(\alpha, \beta, 1)$ for any $\mathcal{T}_+$.

Next, we show that in the class of the GS models, the homogeneity property is satisfied exclusively for models, where the functions ψ^x and ψ^y are of specific forms, and the directions are (G1) directions, in the graph models, or $(\boldsymbol{x}_o, \boldsymbol{0})$ in the input-oriented and $(\boldsymbol{0}, \boldsymbol{y}_o)$ in the output-oriented models.

We start with the following theorem that provides a necessary and sufficient condition for homogeneity of the GS model with (G1) directions. The proof can be found in Appendix A.

Theorem 13. *The GS model with (G1) directions is homogeneous of degree $(\alpha, \beta, 1)$ if and only if*

$$\psi^x(\theta) = \theta^{-\alpha}, \ \alpha \in [-1, 0), \qquad \psi^y(\theta) = \theta^{-\beta}, \ \beta > 0, \quad with \quad \mathrm{dom}(\psi^x) = \mathrm{dom}(\psi^y) = (0, \infty). \quad (21)$$

The next theorem provides a necessary and sufficient condition for the homogeneity of input- and output-oriented models. The proof is analogous to the proof of Theorem 13 and hence is omitted.

Theorem 14. *(a) The input-oriented GS model with $(\boldsymbol{g}^x, \boldsymbol{g}^y) = (\boldsymbol{x}, \boldsymbol{0})$ is homogeneous of degree $(\alpha, 0, 1)$ if and only if $\psi^x(\theta) = \theta^{-\alpha}, \theta > 0, \ \alpha \in [-1, 0)$.*

(b) The output-oriented GS model with $(\boldsymbol{g}^x, \boldsymbol{g}^y) = (\boldsymbol{0}, \boldsymbol{y})$ is homogeneous of degree $(0, \beta, 1)$ if and only if $\psi^y(\theta) = \theta^{-\beta}, \ \theta > 0, \ \beta > 0$.

Finally, the following theorem states that models with the (G2)-(G6) directions discussed in this article are not homogeneous. Its proof is in Appendix A.

Theorem 15. *The GS model with directions $(\boldsymbol{g}_o^x, \boldsymbol{g}_o^y) \gneqq 0$ of the form $(\boldsymbol{g}^x, \boldsymbol{g}^y) = (\mathbf{h}^x, \mathbf{h}^y)$ or $(\boldsymbol{g}^x, \boldsymbol{g}^y) = (\gamma \boldsymbol{x}_o - \mathbf{h}^x, \mathbf{h}^y - \delta \boldsymbol{y}_o)$, where $(\mathbf{h}^x, \mathbf{h}^y)$ is a constant vector, is not homogeneous.*

Remark 13. Theorems 13 and 14 cover the known results for the HDF, BCC-I, and BCC-O models. Necessarily, in these cases one has $\psi^x(\theta) = \theta$ and $\psi^y(\theta) = \theta^{-1}$. Thus, $\alpha = -1$ and $\beta = 1$. This implies that HDF, BCC-I and BCC-O are homogeneous of degrees $(-1, 1, 1)$, $(-1, 0, 1)$, and $(0, 1, 1)$, respectively. These results were originally shown in Färe et al. (1985) and later in Cuesta and Zofío (2005).

The measure defined by $\psi^x(\theta) = \theta, \psi^y(\theta) = \theta^{-p}$, where $p > 0$, is homogeneous of degree $(-1, p, 1)$. Furthermore, the measure of the generalised distance function, defined by $\psi^x(\theta) = \theta^{1-p}, \psi^y(\theta) = \theta^{-p}, p \in [0, 1]$ and introduced in Chavas and Cox (1999) is homogeneous of degree $(p - 1, p, 1)$, see also Proposition 2 in Chavas and Cox (1999).

Remark 14 (g-homogeneity). The standard formulation of DDF-g models is in terms of an inefficiency measure δ. This measure can be obtained from our GS models by the transformation $\delta = 1 - \theta$. It is well known that δ satisfies the so-called *g-homogeneity* (Hudgins and Primont (2007)) of degree $(1, -1)$ with the following meaning: scaling each direction $\boldsymbol{g}$ by $\mu > 0$ results in scaling the measure δ by μ^{-1}.

Let us note that the concept of g-homogeneity of degree $(1, -1)$ could be generalised to more general powers of μ, i.e. to g-homogeneity of degree (p, q). However, a similar procedure as used in the proof of Theorem 13 shows that the GS model is g-homogeneous of degree (p, q) if and only if $p = 1, q = -1$ and $\psi^x = \theta, \psi^y = 2 - \theta$, which corresponds to linear DDF-g models.

5. Extensions

Section 4 has introduced some theoretical tools useful in analysing the ten desirable properties (P1)–(P10). With the exception of (P10), the properties were analysed over arbitrary data. In the

Direction notation	|G1|	|G2|	(G3)	|G4|	(G5)	(G6)
$\boldsymbol{g}_o^x$	$\|\boldsymbol{x}_o\|$	$\|\boldsymbol{x}_o - \boldsymbol{x}^{\min}\|$	$\boldsymbol{x}^{\max} - \boldsymbol{x}^{\min}$	$\|\boldsymbol{x}^{ev}\|$	$\boldsymbol{x}^{sd}$	$\boldsymbol{e}$
$\boldsymbol{g}_o^y$	$\|\boldsymbol{y}_o\|$	$\|\boldsymbol{y}^{\max} - \boldsymbol{y}_o\|$	$\boldsymbol{y}^{\max} - \boldsymbol{y}^{\min}$	$\|\boldsymbol{y}^{ev}\|$	$\boldsymbol{y}^{sd}$	$\boldsymbol{e}$
(P2) indentification	✗(✗)	✗*(✗)	✗(✗)	✗(✗)	✗(✗)	✗(✗)
(P3) strong efficiency	✗(✗)	✗*(✗)	✗(✗)	✗(✗)	✗(✗)	✗(✗)
(P4) $\theta_o^* \in [0,1]$	✓**(✓)	✓(✓)	✓(✓)	✗(✓)	✗(✓)	✗(✓)
(P5) unit invariance	✓(✓)	✓(✓)	✓(✓)	✓(✓)	✓(✓)	✗(✗)
(P6) translation invariance	✗(✗)	✓(✓)	✓(✓)	✗(✗)	✓(✓)	✓(✓)
(P7) monotonicity	✓**(✓)	✓(✓)	✓(✓)	✓(✓)	✓(✓)	✓(✓)
(P8) strict monotonicity	✗(✗)	✗*(✗)	✗(✗)	✗(✗)	✗(✗)	✗(✗)
(P9) super-efficiency	✓*(✓**)	✓*(✗)	✓*(✗)	✓*(✗)	✓*(✗)	✓(✗)
(P10) homogeneity	✗(✓**)	✗(✗)	✗(✗)	✗(✗)	✗(✗)	✗(✗)

Table 3: Properties of the DDF-g model for $\psi^x(\theta) = \theta$, $\psi^y(\theta) = 2 - \theta$ and the HDF-g model for $\psi^x(\theta) = \theta$, $\psi^y(\theta) = \frac{1}{\theta}$ (in brackets) with respect to different choices of vector pairs $\boldsymbol{g}_o^x, \boldsymbol{g}_o^y$. ✓* – the property is satisfied for positive directions; ✓** – the property is satisfied for positive data but not for general data; ✗* – the property holds only for some very specific $\mathcal{T}$.

remarks of Section 4, these tools were applied to standard path-based models over non-negative data. Next, we expand the analysis to examine the first nine properties of standard models on arbitrary data (Section 5.1). Furthermore, we will demonstrate how the theoretical tools of Section 4 facilitate the design of new models, where all or almost all desirable properties are satisfied (Subsection 5.2).

5.1. Properties of standard path-based models over arbitrary data

The particular choices of the vectors $\boldsymbol{g}_o$ in Table 3 are based on the list found in Table 1. Since we now allow for negative data, the non-negativity of directions is ensured by taking absolute values where necessary. This is the case of expressions for (G1) a (G4), but also for (G2), where the values of units outside $\mathcal{T}$ could harm the assumption $\boldsymbol{g}_o \gneqq \boldsymbol{0}$. Modified directions with absolute values are indicated by |G1|, |G2|, and |G4|.

The list of desirable properties in Table 3 omits the property (P1), which in the (GS) scheme is satisfied regardless of the choice of ψ^x, ψ^y or $\boldsymbol{g}_o$. Observe that all the models in Table 3 together fail on the same properties (P2), (P3), and (P8).[10] Since (P2) is not satisfied in any of the models considered, failure of (P3) is inevitable in view of the implication (P3) $\Rightarrow$ (P2) shown in Remark 4.

- **Strong efficiency of projections** (P3). According to Remark 4, all models fail (P3) with non-negative data and therefore also for arbitrary data.

---

[10]It may happen that DDF-(G2) meets (P2), (P3), and (P8) over some very specific $\mathcal{T}$ — see Figure 2.

- **Boundedness** (P4): For $\psi^x(\theta) = \theta$, $\psi^y(\theta) = \frac{1}{\theta}$, allowing arbitrary data does not produce a change; for all directions $\boldsymbol{g}_o \gneqq 0$ one again obtains $\theta_o^* > 0$. Similarly, for $\psi^x(\theta) = \theta$, $\psi^y(\theta) = 2 - \theta$, and directions $|\text{G2}| = (\text{G2})$ and (G3) we have $\theta_o^* \geq 0$ by Theorem 6. However, for directions $|\text{G1}|$, the status changes; negative data may lead to a negative score, as shown in the example of one input, one output with two units $A = (-1, 3), B = (3, 1)$, where the directional vector for B is $(3, 1)$ and its score equals $-1/3$. However, if at least one of the inputs takes only positive values, then Theorem 6 implies that $\theta_o^* > 0$ for all $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T}$. This is in agreement with the results of Kerstens and Van de Woestyne (2011).

  By Remark 5, DDF-g models with directions (P4)–(P6) do not guarantee non-negative efficiency scores for positive data and, therefore, the same is true for general data.

- **Unit and translation invariance** (P5) **and** (P6)**.** There is no change compared to the positive data case and the conclusions of Remarks 6, 7 remain valid.

- **Monotonicity** (P7)**.** Allowing negative data does not affect the status of (P7) for directions that do not explicitly depend on $(\boldsymbol{x}_o, \boldsymbol{y}_o)$. Therefore, the models with directions (G3), $|\text{G4}|$, (G5), and (G6) satisfy (P7). Likewise, the directions $|\text{G2}| = (\text{G2})$ meet the assumptions of Theorem 9, and therefore here there is also no change.

  The situation is different for DDF with the directions $|\text{G1}|$, which do not satisfy the assumptions of Theorem 9. To see that (P7) may fail, consider a one-input, one-output counterexample with three units: efficient unit $A = (-1, \ 4)$ and two inefficient units $B = (0.1, \ 0)$, $C = (1, \ 0)$. Although B dominates C, one has $\theta_B^* = -10 < \theta_C^* = -1$. We remark that directions $|\boldsymbol{x}_o|$, resp. $|\boldsymbol{y}_o|$ appear in Cheng et al. (2013) in connection with radial oriented models. There, it is observed that monotonicity may fail if zero is an internal point of the data.

- **Strict monotonicity** (P8)**.** According to Remark 9, all models fail (P8) with non-negative data and therefore also for arbitrary data.

- **Super-efficiency** (P9)**.** DDF-g models with directions $|\text{G1}|$, $|\text{G2}|$, (G3), $|\text{G4}|$, (G5), and (G6) retain the property (P9) for arbitrary data (as long as the directions are positive). Similarly, the negative results for HDF-g models with the directions $|\text{G2}|$, (G3), $|\text{G4}|$, (G5), and (G6) carry over to arbitrary data. Theorem 12 does not resolve the status of HDF-g models with $|\text{G1}|$ directions. However, one can find an example where the model is infeasible. Consider a single-input, single-output case, where technology $\mathcal{T}$ is generated by a single unit $A = (2, 0)$. For $B = (1, 10) \notin \mathcal{T}$, whose $|\text{G1}|$ direction reads $\boldsymbol{g}_B = (1, 10)$, the output inequality $0 \geq 10\,\theta^{-1}$ is not satisfied for any $\theta > 0$.

- **Homogeneity** (P10)**.** Theoretical tools are derived only for positive data in Section 4.9. Nevertheless, it can be shown that the statement of Theorem 15 can be extended to all GS models over negative data.
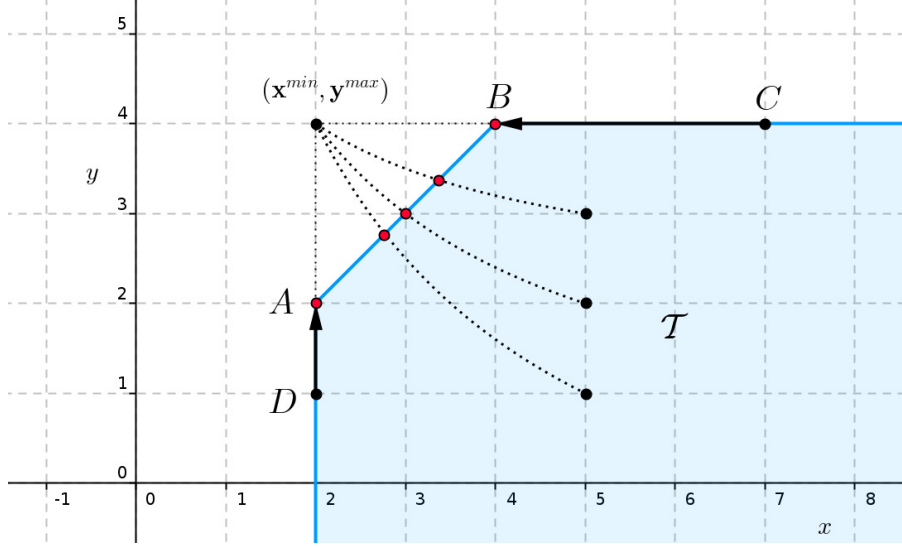
Figure 4: All paths with directions (G2.0) intersect at the point $(\boldsymbol{x}^{\min}, \boldsymbol{y}^{\max})$ (Lemma 4). Furthermore, (G2.0) projects all units in two-dimensional technology $\mathcal{T}$ onto the strongly efficient frontier.

5.2. Non-standard directions with good properties

We will focus our attention on the (G2) directions, which we have already shown produce the invariance properties (P5) and (P6) for arbitrary ψ^x and ψ^y satisfying (A1)–(A4) (see Remarks 6 and 7). These directions allow for a small modification that will also ensure the validity of (P5) and (P7). The idea is to choose $\boldsymbol{g}_o$ so that the corresponding path $\boldsymbol{\phi}_o$ crossing $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ at $\theta = 1$ passes through the point $(\boldsymbol{x}^{\min}, \boldsymbol{y}^{\max})$ at $\theta = \theta_{\min}$ (see Figure 3). Then $\theta_{\min}$ serves as the lower bound for the efficiency score, as indicated in the next lemma. Its proof follows from a simple calculation and is therefore omitted.

Lemma 4. *Let $\theta_{\min} \in [0, 1) \cap \mathcal{D}$ and $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{T} \setminus \{(\boldsymbol{x}^{\min}, \boldsymbol{y}^{\max})\}$. The path $\boldsymbol{\phi}_o(\theta)$ runs through $(\boldsymbol{x}^{\min}, \boldsymbol{y}^{\max})$ at $\theta = \theta_{\min}$, i.e., $\boldsymbol{\phi}_o(\theta_{\min}) = (\boldsymbol{x}^{\min}, \boldsymbol{y}^{\max})$ if and only if the directions $\boldsymbol{g}_o^x, \boldsymbol{g}_o^y$ satisfy*

$$\boldsymbol{g}_o^x = \frac{\boldsymbol{x}_o - \boldsymbol{x}^{\min}}{1 - \psi^x(\theta_{\min})}, \quad \boldsymbol{g}_o^y = \frac{\boldsymbol{y}^{\max} - \boldsymbol{y}_o}{\psi^y(\theta_{\min}) - 1}. \tag{G2.0}$$

Remark 15. The GS models with directions (G2.0) satisfy (P5) and (P6) for arbitrary ψ^x and ψ^y with properties (A1)–(A4) by Theorems 7 and 8. By Lemma 4, the same holds for the boundedness property (P4). Finally, after substituting directions (G2.0) into the right-hand sides in (12) and performing simple calculations, the monotonicity of $\boldsymbol{\phi}(\theta)$ obtains for each $\theta \geq \theta_{\min}$. Hence, the monotonicity property (P7), too, holds universally with directions (G2.0).

Regarding property (P9), the situation is more nuanced for two reasons. First, the super-efficiency considers all points $(\boldsymbol{x}_o, \boldsymbol{y}_o) \notin \mathcal{T}$, of which only a small subset with $\boldsymbol{x}_o \geq \boldsymbol{x}^{\min}$ and $\boldsymbol{y}_o \leq \boldsymbol{y}^{\max}$ yields non-negative directions (G2.0). This leads us to consider the absolute values of the expressions in (G2.0) as explained in Subsection 5.1. Second, if the range of ψ^y is bounded from below, then the feasibility is compromised on the output side for units (outside $\mathcal{T}$) with very high

ψ^y	$2-\theta$	$\theta^{-p}, p>0$	$1-\ln\theta$	$\mathrm{e}^{1-\theta}$
domain	$(-\infty,\infty)$	$(0,\infty)$	$(0,\infty)$	$(-\infty,\infty)$
image	$(-\infty,\infty)$	$(0,\infty)$	$(-\infty,\infty)$	$(0,\infty)$
$\theta_{\min}$	0	$1/2$	e^{-1}	0
$\boldsymbol{g}_o^x$	$|\boldsymbol{x}_o-\boldsymbol{x}^{\min}|$	$2|\boldsymbol{x}_o-\boldsymbol{x}^{\min}|$	$\frac{\mathrm{e}}{\mathrm{e}-1}|\boldsymbol{x}_o-\boldsymbol{x}^{\min}|$	$|\boldsymbol{x}_o-\boldsymbol{x}^{\min}|$
$\boldsymbol{g}_o^y$	$|\boldsymbol{y}^{\max}-\boldsymbol{y}_o|$	$\frac{1}{2^p-1}|\boldsymbol{y}^{\max}-\boldsymbol{y}_o|$	$|\boldsymbol{y}^{\max}-\boldsymbol{y}_o|$	$\frac{1}{\mathrm{e}-1}|\boldsymbol{y}^{\max}-\boldsymbol{y}_o|$
(P1), (P4)–(P7)	✓	✓	✓	✓
(P2), (P3), (P8)	✗*	✗*	✗*	✗*
(P9)	✓*	✗	✓*	✗

Table 4: Values and properties of directions (G2.0) taken in absolute value for $\psi^x(\theta)=\theta$ and three choices of of ψ^y and $\theta_{\min}$, ✓* - the property is satisfied for positive directions; ✗* – the property holds only for some very specific $\mathcal{T}$.

outputs. Thus, for ψ^y with a bounded range, the super-efficiency fails. In contrast, if the range of ψ^y is unbounded, by Theorem 12(b), a GS model satisfies super-efficiency for those directions (G2.0) that are positive, i.e., super-efficiency holds for all units outside $\mathcal{T}$ except for a negligible set where some direction components are zero.

With respect to properties (P2), (P3), and (P8), the situation is similar to the case of the DDF-(G2) model. The existence of a point $(\boldsymbol{x}_o,\boldsymbol{y}_o)\in\partial^W\mathcal{T}$ with $\theta_o^*=1$ will cause all three properties to fail simultaneously (for (P2) and (P3) this is immediate; for (P8) see Theorem 10). Figure 3 shows that such points can also be found for directions (G2.0) in technological sets of dimension greater than 2. In the case of two-dimensional $\mathcal{T}$, the properties can be satisfied as illustrated in Figure 2 for DDF-(G2) and in Figure 4 for linear $\psi^x=\theta$ and hyperbolic $\psi^y=\frac{1}{\theta}$.

We now examine the (GS) scheme for linear $\psi^x(\theta)=\theta$ and four choices of ψ^y that satisfy the assumptions (A1)–(A4): $\psi_1^y(\theta)=2-\theta$; $\psi_2^y(\theta)=\theta^{-p}$, $p>0$; $\psi_3^y(\theta)=1-\ln\theta$; and $\psi_4^y(\theta)=\mathrm{e}^{1-\theta}$. These functions offer a variety of combinations of domains and ranges, as shown in Table 4. Observe that ψ_1^y and ψ_2^y with $p=1$, have already appeared in the context of the DDF-g and HDF-g models, respectively. To evaluate directions (G2.0), one has to specify the value of $\theta_{\min}\in[0,1)$. Only for ψ_1^y and ψ_4^y can one choose the ideal value $\theta_{\min}=0$. The remaining two cases do not have zero in their domain, and for those, we select ad hoc positive values of $\theta_{\min}$ that yield simple expressions for directions. Intuitively, the choice of $\theta_{\min}$ affects the "scaling" of optimal scores for non-efficient units. The resulting values are summarised in Table 4. Observe that the directions (G2.0) for ψ_1^y with the choice of $\theta_{\min}=0$ lead to (G2). On the contrary, ψ_2^y with $p=1$ and $\theta_{\min}=\frac{1}{2}$ generate directions whose x component is multiplied by a factor of 2 compared to (G2).

The statuses with respect to (P1)–(P9) follow from Remark 15, where they were analysed for general ψ^x and ψ^y. Note that the models with ψ_2^y and ψ_4^y fail (P9) due to the boundedness of their ranges. The property (P10) is not included in Table 4, since the results summarised there correspond to (G2.0) directions for which the models are not homogeneous; see Theorem 15.

5.3. Numerical illustration with mixed data

We present a numerical example that examines the directions (G1) and (G2.0) for the models with $\psi^x(\theta) = \theta$ and the functions $\psi^y(\theta)$ appearing in Table 4. We demonstrate the applicability of our approach to a data-set containing negative data. To evaluate the efficiency score, as well as to obtain the second phase result, we use the CVX modelling system (see Grant and Boyd (2014), Grant and Boyd (2008)) to solve convex programmes, implemented in Matlab. Data for 30 Taiwanese electrical machinery producers are taken from Table 5 of Tone et al. (2020) and consist of two inputs and two outputs. The minimal and maximal values of the data sample can be seen in Table 5. The results are presented in Tables B.6 and B.7 in Appendix B.

	Inputs		Outputs	
	Cost of sales	R&D expenses	Net income	Return
Variable	x_1	x_2	y_1	y_2
min	418,242	1,319	-561,965	-0.2583
max	16,655,569	634,436	3,092,358	3.53714

Table 5: Minimal and maximal values of the data sample used in the numerical example.

As stated in Section 5.1, the DDF-g model with direction |G1| may lead to negative efficiency score, in general. However, this happens only if a negative value occurs in each of the inputs in the data-sample. It can be seen from Table 5 that the inputs are positive and therefore the efficiency score is bounded bellow by zero.

The choice of a particular function ψ^y has only minimal influence on the ranking of the best-placed units, but the effect is more pronounced for very inefficient units, where the change in ranking can reach 10 places out of 30 companies. On the other hand, the choice of ψ^y has little impact on the number of units projected onto the strongly efficient frontier, where the main determinant is the choice of directions. The number of strongly efficient projections fluctuates between 12 and 13 for the (G1) directions, while it is between 20 and 23 units for the (G2.0) directions. The choice of the scaled (G2) direction (G2.0) not only leads to a larger number of correct efficiency evaluations, but in the case of the hyperbolic and logarithmic model, it also improves the discriminatory power of the model.

6. Conclusions

The paper formulates a general scheme for path-based models that includes BCC, DDF, and HDF as special cases. The models belonging to the scheme are analysed in light of ten desirable properties. Only the unique projection property (P1) is universal; the remaining properties must be examined individually and generally depend on the choice of directions, functions ψ, and possibly also on the *configuration of data* that defines the technological set $\mathcal{T}$.

The article provides mathematical tools (sufficient and/or necessary conditions to verify specific properties) that allow a detailed analysis of individual models within the GS framework. The

usefulness of these tools has been demonstrated on the standard path-based models, on their modifications to general (negative) data, and on models with non-linear and non-hyperbolic functions ψ.

The general scheme is flexible enough to allow modification of existing models designed for positive data to arbitrary (negative) data. For this modification, it is important to rewrite models such as BCC-I (3), BCC-O (4), or HDF (6), in the form of the GS scheme (Table 2), where the expressions for the directions are specified. In the presence of negative data, these expressions are taken as absolute values. This approach is illustrated with a mixed-data numerical example in Section 5.2.

The article provides a complete analysis of the *standard* path-based models and their modifications to negative data with respect to the ten properties (Table 3). In addition to the unique projection property (P1), all standard models satisfy monotonicity (P7). On the other hand, all models fail (P2), (P3), and (P8) simultaneously for *some* $\mathcal{T}$. Overall, models with (G2) directions satisfy or partially satisfy the highest number of properties, followed by models with (G3) directions.

We have obtained interesting results regarding homogeneity in the general path-based scheme. It is known that in the case of CRS all models in the general scheme meet the property of homogeneity, and thus it could appear that homogeneity and lack of identification are somehow linked. However, the results of this article indicate that this is not the case when the type of technology changes. For VRS, the homogeneity property is satisfied only in a narrow subclass of GS models: In the case of oriented models, homogeneous models are equivalent to BCC-I and BCC-O (Theorem 14), and in the case of graph models, only models with directions (G1) and certain special nonlinearities are homogeneous (Theorems 13 and 15).

Directions have a strong influence on model properties. Models with positive directions fail (P2), (P3), and (P8) for every technology set (Theorems 3 and 11 and Remark 4). This is especially true for positive data and for all considered directions, except (G2). For models with (G2) directions, the status of properties (P2), (P3), and (P8) is ambiguous insofar as it depends on the data configuration. There are examples where all three properties fail, but also data configurations (in particular, single-input, single-output technological sets) where all three are satisfied.

The lack of universal validity of (P2), (P3), and (P8) can be considered a characteristic feature of path-based models. In contrast, Halická and Trnovská (2021) show that the slacks-based models do satisfy these properties. This observation provides a further rationale for distinguishing the two categories of models set out in Russell and Schworm (2018) – the path-based versus slacks-based models. We remark that the classification is able to resolve name paradoxes in the literature, where some models are currently presented with the adjective 'directional' (e.g., Sahoo et al., 2014) suggesting a path-based approach, despite being in reality slacks-based.

The paper opens up several new research directions.

– The simultaneous failure of (P2), (P3), and (P8) raises the question of whether these properties are linked. It is fairly immediate that (P3) implies (P2); such implication is valid also outside our scheme. However, the connection between (P3) and (P8) is not known or obvious.

We aim to address this problem in future work.

– The general scheme (introduced here in an envelopment form) allows the derivation of a dual multiplicative model, which can then be related to the shadow profit inefficiency. This opens the way for obtaining relationships between technical and overall profit inefficiencies similar to those known for DDF (Chambers et al., 1998).

– To simplify the exposition, we have only considered technologies with variable returns to scale. The VRS assumption is essential to allow for negative data (see, e.g., Portela et al., 2004). On the other hand, many results in the paper can be adapted to technology sets with other types of returns to scale, at the cost of considering only non-negative data.

– The numerical results in Section 5.3 indicate that directions (G2.0) may be helpful in practical applications in the sense that a higher number of DMUs are projected onto the strongly efficient frontier compared to other directions. It is natural to ask whether this finding holds up for other data sets.

Acknowledgements

Appendix A. Proofs of selected theorems

Proof of Theorem 1. (a) This follows from the assumptions placed on $\boldsymbol{g}_o$ in (16), the monotonicity properties of ψ^x, ψ^y and the definition of $\boldsymbol{\phi}_o$ in (12).

(b) This follows from (12) and (A3).

(c) This part follows from (12), (A1), and (A4).

(d) The assumptions on $\boldsymbol{g}_o$ and part (c) of this theorem yield $\lim_{\theta \to a_+^x} \phi_{io}^x(\theta) = -\infty$ for some i or $\lim_{\theta \to a_+^y} \phi_{ro}^y(\theta) = +\infty$ for some r . From this, part (b) of this theorem, and the continuity of $\boldsymbol{\phi}_o$, it follows that there exists a finite $\bar{\theta}$ such that $\phi_{ro}^y(\bar{\theta}) > y_r^{\max}$ or $\phi_{io}^x(\bar{\theta}) < x_i^{\min}$ (see Figure 1). Therefore, by Lemma 1(a) one obtains $\boldsymbol{\phi}_o(\bar{\theta}) \notin \mathcal{T}$. The property (a) of this theorem then shows that for each $\bar{\theta}$ such that $\boldsymbol{\phi}_o(\bar{\theta}) \notin \mathcal{T}$ one has $\boldsymbol{\phi}_o(\theta) \notin \mathcal{T}$ for all $\theta \leq \bar{\theta}$.

On the other hand, $\phi_o(1) = (x_o, y_o) \in \mathcal{T}$ and the property (a) of this theorem implies that if $\phi_o(\hat{\theta}) \in \mathcal{T}$ for some $\hat{\theta}$, then $\phi_o(\theta) \in \mathcal{T}$ for all $\theta \geq \hat{\theta}$. Now, by the continuity of $\boldsymbol{\phi}_o$, there exists a finite θ_o^* such that $\theta_o^* = \sup\{\theta : \boldsymbol{\phi}_o(\theta) \notin \mathcal{T}\} = \inf\{\theta : \boldsymbol{\phi}_o(\theta) \in \mathcal{T}\}$. We show that $\boldsymbol{\phi}_o(\theta_o^*)$ belongs to the boundary $\partial \mathcal{T}$ of the closed set $\mathcal{T}$. Assume, in contradiction, that $\boldsymbol{\phi}_o(\theta_o^*) \in \text{int}\mathcal{T}$. From the

continuity of $\boldsymbol{\phi}_o$ it follows that $\boldsymbol{\phi}_o(\theta) \in \text{int}\mathcal{T}$ for all θ close enough to θ_o^*, and hence also for some $\theta < \theta_o^*$ that contradicts $\theta_o^* = \inf\{\theta : \boldsymbol{\phi}_o(\theta) \in \mathcal{T}\}$. $\qquad\square$

Proof of Theorem 4. Assume by contradiction that $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ is strongly efficient and that there exists an optimal solution $(\boldsymbol{\lambda}_o^*, \theta_o^*)$ of $(GS)_o$ such that at least one component of the corresponding optimal slacks is positive. Then the unit $(\boldsymbol{X}\boldsymbol{\lambda}_o^*, \boldsymbol{Y}\boldsymbol{\lambda}_o^*)$ differs from $(\hat{\boldsymbol{x}}_o, \hat{\boldsymbol{y}}_o)$, belongs to $\mathcal{T}$ and dominates $(\hat{\boldsymbol{x}}_o, \hat{\boldsymbol{y}}_o)$, and therefore $(\hat{\boldsymbol{x}}_o, \hat{\boldsymbol{y}}_o)$ is not strongly efficient.

To prove the reverse implication, we assume that $(\hat{\boldsymbol{x}}_o, \hat{\boldsymbol{y}}_o) \in \partial^W \mathcal{T}$. According to Lemma 1(d) there exists $(\boldsymbol{d}^x, \boldsymbol{d}^y) \gneqq \boldsymbol{0}$, $(\boldsymbol{d}^x, \boldsymbol{d}^y) \not\gtrless \boldsymbol{0}$ such that $(\hat{\boldsymbol{x}}_o - \boldsymbol{d}^x, \hat{\boldsymbol{y}}_o + \boldsymbol{d}^y) \in \mathcal{T}$. By the definition of $\mathcal{T}$, there then exists $\hat{\boldsymbol{\lambda}} \geq \boldsymbol{0}$ such that $\boldsymbol{e}^T\hat{\boldsymbol{\lambda}} = 1$, $\boldsymbol{X}\hat{\boldsymbol{\lambda}} \leq \hat{\boldsymbol{x}}_o - \boldsymbol{d}^x$, $\boldsymbol{Y}\hat{\boldsymbol{\lambda}} \geq \hat{\boldsymbol{y}}_o + \boldsymbol{d}^y$. This implies that $(\hat{\boldsymbol{\lambda}}, \theta_o^*)$ is an optimal solution to $(GS)_o$ for $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ and at least one component of inequality (11b) or (11c) is strict. $\qquad\square$

Proof of Theorem 7. The input inequality (11b) for transformed inputs reads

$$C\sum_{j=1}^{n}\boldsymbol{x}_j\lambda_j = \sum_{j=1}^{n}(C\boldsymbol{x}_j)\lambda_j \leq C\boldsymbol{x}_o + (\psi^x(\theta) - 1)\boldsymbol{g}_o^{Cx} \Leftrightarrow C\sum_{j=1}^{n}\boldsymbol{x}_j\lambda_j \leq C(\boldsymbol{x}_o + (\psi^y(\theta) - 1)\boldsymbol{g}_o^x).$$

Since C is diagonal with positive entries, this is equivalent to the input inequality (11b) for the original inputs. This and a similar equivalence for outputs imply that conditions (11b)–(11d) before the data transformation are equivalent to those after transformation, and hence the optimal score provided by the model is unchanged. $\qquad\square$

Proof of Theorem 8. Let $\boldsymbol{c} \in \mathbb{R}^m$ and $\boldsymbol{b} \in \mathbb{R}^s$ are the translation vectors for inputs and outputs, and hence the translated inputs and outputs are $\boldsymbol{x}_j + \boldsymbol{c}$ and $\boldsymbol{y}_j + \boldsymbol{b}$, for all $j = 1, \ldots, n$ respectively. The condition $\sum_{j=1}^{n}\lambda_j = 1$ and the invariance of $\boldsymbol{g}_o^x$ on the translation of inputs imply that the input inequality (11b) of $(GS)_o$ for translated inputs is equivalent to the input inequality (11b) of $(GS)_o$ for original inputs:

$$\sum_{j=1}^{n}(\boldsymbol{x}_j + \boldsymbol{c})\lambda_j = \sum_{j=1}^{n}\boldsymbol{x}_j\lambda_j + \boldsymbol{c} \leq \boldsymbol{x}_o + \boldsymbol{c} + (\psi^x(\theta) - 1)\boldsymbol{g}_o^x \Leftrightarrow \sum_{j=1}^{n}\boldsymbol{x}_j\lambda_j \leq \boldsymbol{x}_o + (\psi^x(\theta) - 1)\boldsymbol{g}_o^x.$$

Similar equivalence in outputs proves the theorem. $\qquad\square$

Proof of Theorem 12. (a) The proof runs along the lines of the proof of Theorem 1(d). The fact that $\theta_o^* > 1$ is a simple consequence of the strict monotonicity of $\phi_o^x(\theta)$, $\phi_o^y(\theta)$, $(\boldsymbol{x}_o, \boldsymbol{y}_o) \notin \mathcal{T}$, and $\mathcal{T}$ being closed.

(b) Since $\boldsymbol{X}\boldsymbol{\lambda} \leq \boldsymbol{x}^{\max}$ and $\boldsymbol{Y}\boldsymbol{\lambda} \geq \boldsymbol{y}^{\min}$, a sufficient condition for the feasibility of $(GS)_o$ is the existence of $\bar{\theta} \in \mathcal{D}$, such that

$$\boldsymbol{x}^{\max} - \boldsymbol{x}_o + \boldsymbol{g}_o^x \leq \psi^x(\bar{\theta})\boldsymbol{g}_o^x, \tag{A.1}$$

$$\boldsymbol{y}^{\min} - \boldsymbol{y}_o + \boldsymbol{g}_o^y \geq \psi^y(\bar{\theta})\boldsymbol{g}_o^y. \tag{A.2}$$

Since $\boldsymbol{g}_o^x > 0$, ψ^x is increasing and $\lim_{\theta \to +\infty} \psi^x(\theta) = +\infty$, the constraint (A.1) is satisfied for $\bar{\theta}$ sufficiently large. Since ψ^y is decreasing, the case $\lim_{\theta \to +\infty} \psi^x(\theta) = +\infty$ implies that the constraint

(A.2) is satisfied for $\bar{\theta}$ sufficiently large. On the other hand, if $\lim_{\theta \to \infty} \psi^x(\theta) = 0$, then the assumptions $\boldsymbol{y}^{\min} - \boldsymbol{y}_o + \boldsymbol{g}_o^y > 0$ and $\boldsymbol{g}_o^y > 0$ guarantee that (A.2) is satisfied for $\bar{\theta}$ sufficiently large.

(c) Let us first consider the case (i). By assumption, for each feasible $\boldsymbol{\lambda}$ ($\boldsymbol{\lambda} \geq 0, \boldsymbol{e}^T\boldsymbol{\lambda} = 1$), there exists r such that $(\boldsymbol{Y\lambda})_r < y_{ro} - g_{ro}^y$. The assumption on ψ^y implies $(\boldsymbol{Y\lambda})_r < y_{ro} - g_{ro}^y < y_{ro} - g_{ro}^y + \psi^y(\theta)g_o^y$ for all $\theta \in \mathrm{dom}(\psi^y)$. Hence we have shown that $\boldsymbol{Y\lambda} \geq \boldsymbol{y}_o - \boldsymbol{g}_o^y + \psi^y(\theta)\boldsymbol{g}_o^y$ is violated for all feasible $\boldsymbol{\lambda}$ and $\theta \in (1, \infty)$ and therefore $(GS)_o$ is infeasible. Consider now the case (ii). Since $(\boldsymbol{Y\lambda})_r \leq y_r^{\max}$ for all feasible $\boldsymbol{\lambda}$, we have $(\boldsymbol{Y\lambda})_r < y_{ro} - g_{ro}^y$ and the claim follows by part (i). $\qquad \square$

Proof of Theorem 13. First assume that GS model with (G1) directions is homogeneous of degree $(\alpha, \beta, 1)$, i.e. for any technology set $\mathcal{T}_+$ the corresponding efficiency measure satisfies $\theta^*(\mu^\alpha \boldsymbol{x}, \mu^\beta \boldsymbol{y}) = \mu\theta^*(\boldsymbol{x}, \boldsymbol{y})$ for each $(x, y) \in \mathcal{T}_+$ and each $\mu > 0$. Let $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathbb{R}_+^m \times \mathbb{R}_+^s$ and let $\mathcal{T}_+$ be arbitrary, such that $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \partial^S \mathcal{T}_+$. Then any efficiency measure defined over $\mathcal{T}_+$ satisfies $\theta^*(\mu^\alpha \boldsymbol{x}_o, \mu^\beta \boldsymbol{y}_o) = \mu$ for all $\mu > 0$ and the projection point for $(\mu^\alpha \boldsymbol{x}_o, \mu^\beta \boldsymbol{y}_o)$ is given by

$$\boldsymbol{\phi}^x(\theta^*(\mu^\alpha \boldsymbol{x}_o, \mu^\beta \boldsymbol{y}_o)) = \boldsymbol{\phi}^x(\mu) = \psi^x(\mu)\mu^\alpha \boldsymbol{x}_o, \quad \boldsymbol{\phi}^y(\theta^*(\mu^\alpha \boldsymbol{x}_o, \mu^\beta \boldsymbol{y}_o)) = \boldsymbol{\phi}^y(\mu) = \psi^y(\mu)\mu^\beta \boldsymbol{y}_o.$$

The set $\mathcal{P} := \{(\boldsymbol{\phi}^x(\mu), \boldsymbol{\phi}^y(\mu)), \ \mu > 0\}$ is a smooth path satisfying $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{P} \subset \partial \mathcal{T}_+$. Recall that $\mathcal{T}_+$ was an arbitrary technology set satisfying $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \partial^S \mathcal{T}_+$. Therefore it must hold $(\boldsymbol{\phi}^x(\mu), \boldsymbol{\phi}^y(\mu)) = (\boldsymbol{x}_o, \boldsymbol{y}_o), \forall \mu > 0$, which is equivalent to

$$\psi^x(\mu)\mu^\alpha = 1, \quad \psi^y(\mu)\mu^\beta = 1, \quad \forall \mu > 0.$$

Therefore $\psi^x(\theta) = \theta^{-\alpha}, \psi^y(\theta) = \theta^{-\beta}$; the ranges for α, β follow from (A2).

Now assume that the functions ψ^x and ψ^y are of the form (21). It is easy to see that these functions satisfy both the assumptions (A1)–(A4) and the homogeneity properties:

$$\psi^x(\theta/\mu) = \mu^\alpha \psi^x(\theta), \quad \psi^y(\theta/\mu) = \mu^\beta \psi^y(\theta). \tag{A.3}$$

From (A.3) it follows that $\psi^x(\mu)\mu^\alpha = 1, \psi^y(\mu)\mu^\beta = 1$ and therefore the efficiency measure

$$\theta^*(\mu^\alpha \boldsymbol{x}, \mu^\beta \boldsymbol{y}) = \min\{\theta \mid (\psi^x(\theta)\mu^\alpha \boldsymbol{x}, \psi^y(\theta)\mu^\beta \boldsymbol{y}) \in \mathcal{T}\}$$

is well-defined provided $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{T}_+$, since the corresponding minimisation problem is feasible and the optimal value is attained (Theorem 1). Using (A.3) and substitution $\tilde{\theta} := \theta/\mu$ we get

$$\min\{\theta \mid (\psi^x(\theta)\mu^\alpha \boldsymbol{x}, \psi^y(\theta)\mu^\beta \boldsymbol{y}) \in \mathcal{T}\} = \mu \min\{\theta/\mu \mid (\psi^x(\theta/\mu)\boldsymbol{x}, \psi^y(\theta/\mu)\boldsymbol{y}) \in \mathcal{T}\}$$
$$= \mu \min\{\tilde{\theta} \mid (\psi^x(\tilde{\theta})\boldsymbol{x}, \psi^y(\tilde{\theta})\boldsymbol{y}) \in \mathcal{T}\} = \mu\theta^*(\boldsymbol{x}, \boldsymbol{y})$$

which proves the homogeneity of θ^*. $\qquad \square$

Proof of Theorem 15. Consider a GS model with directions $(\boldsymbol{g}^x, \boldsymbol{g}^y) = (\mathbf{h}^x, \mathbf{h}^y)$, where $(\mathbf{h}^x, \mathbf{h}^y)$ is a constant vector. Assume by contradiction, that the model is homogeneous of degree $(\alpha, \beta, 1)$. Similarly, as in the proof of Theorem 13 it can be shown that the path $\mathcal{P} = \{(\boldsymbol{\phi}^x(\mu), \boldsymbol{\phi}^y(\mu)), \ \mu > 0\}$ defined by

$$\boldsymbol{\phi}^x(\mu) = \mu^\alpha \boldsymbol{x}_o + (\psi^x(\mu) - 1)\mathbf{h}^x, \quad \boldsymbol{\phi}^y(\mu) = \mu^\beta \boldsymbol{y}_o + (\psi^y(\mu) - 1)\mathbf{h}^y,$$

satisfies $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \mathcal{P} \subset \partial \mathcal{T}_+$ for any $\mathcal{T}_+$, such that $(\boldsymbol{x}_o, \boldsymbol{y}_o) \in \partial^S \mathcal{T}_+$ and hence $(\boldsymbol{\phi}^x(\mu), \boldsymbol{\phi}^y(\mu)) =$

$(\boldsymbol{x}_o, \boldsymbol{y}_o), \forall \mu > 0$. Let $\bar{\mu} > 0$ be arbitrary and fixed. Then

$$(\bar{\mu}^\alpha - 1)\boldsymbol{x}_o = (1 - \psi^x(\bar{\mu}))\mathbf{h}^x, \quad (\bar{\mu}^\beta - 1)\boldsymbol{y}_o = (1 - \psi^y(\bar{\mu}))\mathbf{h}^y.$$

While on the right-hand side the vectors are constant, the left-hand side may vary for different $(\boldsymbol{x}_o, \boldsymbol{y}_o)$, which is a contradiction.

The proof for $(\boldsymbol{g}^x, \boldsymbol{g}^y) = (\gamma \boldsymbol{x}_o - \mathbf{h}^x, \mathbf{h}^y - \delta \boldsymbol{y}_o)$ is analogous.

$\square$

Appendix B. Tables for the numerical illustrations in Section 5.3

Function	$\psi^y(\theta) = 2 - \theta$				$\psi^y(\theta) = 1/\theta$			
Directions	(G1)		(G2.0) $\theta_{\min} = 0$		(G1)		(G2.0) $\theta_{\min} = 0.01$	
DMU	Score	Rank	Score	Rank	Score	Rank	Score	Rank
1	0.66*	15	0.759	28	0.713*	14	0.401	14
2	1*	1-9	1*	1-9	1*	1-9	1*	1-9
3	0.683	14	0.888	12	0.691	15	0.578*	11
4	1*	1-9	1*	1-9	1*	1-9	1*	1-9
5	1*	1-9	1*	1-9	1*	1-9	1*	1-9
6	0.175	29	0.784*	26	0.256	29	0.145*	28
7	0.69	13	0.86*	17	0.763	13	0.254*	19
8	0.283	26	0.801	24	0.3	28	0.167*	24
9	0.126*	30	0.744*	29	0.194*	30	0.112*	30
10	0.372	22	0.777	27	0.526	19	0.156*	27
11	0.37	23	0.707*	30	0.512	20	0.169*	23
12	1*	1-9	1*	1-9	1*	1-9	1*	1-9
13	0.927	10	0.968*	10	0.929	10	0.876	10
14	1*	1-9	1*	1-9	1*	1-9	1*	1-9
15	0.217	28	0.793*	25	0.339	25	0.114	29
16	0.658	16	0.881	15	0.68	16	0.309	16
17	0.22	27	0.807*	22	0.319	27	0.165*	25
18	0.438	19	0.861	16	0.494	21	0.233*	21
19	1*	1-9	1*	1-9	1*	1-9	1*	1-9
20	0.415	21	0.841	18	0.448	22	0.165	26
21	0.335	24	0.804	23	0.408	24	0.27	17
22	1*	1-9	1*	1-9	1*	1-9	1*	1-9
23	1*	1-9	1*	1-9	1*	1-9	1*	1-9
24	0.765	12	0.882	14	0.765	12	0.547	13
25	0.431	20	0.814*	20	0.442	23	0.372	15
26	1*	1-9	1*	1-9	1*	1-9	1*	1-9
27	0.305	25	0.815*	19	0.33	26	0.255	18
28	0.447	18	0.812*	21	0.581	18	0.176*	22
29	0.835*	11	0.933*	11	0.85*	11	0.576*	12
30	0.452	17	0.883*	13	0.598*	17	0.239*	20
Average	0.627		0.880		0.671		0.509	
Minimum	0.126		0.707		0.194		0.112	
Correct	12		21		13		21	

Table B.6: Efficiency scores for models with $\psi^x = \theta$, two choices of ψ^y, directions (G1) and (G2.0). An asterisk next to the efficiency value indicates projection onto the strongly efficient frontier. The last row indicates number of units projected onto the strongly efficient frontier.

Function	$\psi^y(\theta) = \exp(1-\theta)$				$\psi^y(\theta) = 1 - \ln\theta$			
Directions	(G1)		(G2.0) $\theta_{\min} = 0$		(G1)		(G2.0) $\theta_{\min} = 0.01$	
DMU	Score	Rank	Score	Rank	Score	Rank	Score	Rank
1	0,686*	14	0,697	28	0,69*	14	0,565*	18
2	1*	1-9	1*	1	1*	1	1*	1
3	0,686	15	0,833	12	0,687	15	0,686*	12
4	1*	1-9	1*	2	1*	2	1*	2
5	1*	1-9	1*	3	1*	3	1*	3
6	0,19	29	0,698*	26	0,203	29	0,454*	27
7	0,73	13	0,799*	16	0,733	13	0,627*	15
8	0,283	26	0,711	23	0,283	26	0,437	28
9	0,131*	30	0,652*	29	0,137*	30	0,407*	30
10	0,439	20	0,698	27	0,457	20	0,5*	23
11	0,429	22	0,626*	30	0,446	21	0,422	29
12	1*	1-9	1*	4	1*	4	1*	4
13	0,928	10	0,953	10	0,928	10	0,922	10
14	1*	1-9	1*	5	1*	5	1*	5
15	0,232	28	0,708*	25	0,246	28	0,464*	26
16	0,668	16	0,827	13	0,669	16	0,678	13
17	0,238	27	0,726*	22	0,251	27	0,488*	24
18	0,457	19	0,79	17	0,463	19	0,566	17
19	1*	1-9	1*	6	1*	6	1*	6
20	0,425	23	0,763	18	0,428	23	0,512	22
21	0,357	24	0,711	24	0,366	24	0,527*	21
22	1*	1-9	1*	7	1*	7	1*	7
23	1*	1-9	1*	8	1*	8	1*	8
24	0,765	12	0,82	15	0,765	12	0,618	16
25	0,432	21	0,729*	21	0,434	22	0,543*	20
26	1*	1-9	1*	9	1*	9	1*	9
27	0,308	25	0,733*	20	0,31	25	0,484*	25
28	0,51	18	0,741*	19	0,523	18	0,551*	19
29	0,842*	11	0,899*	11	0,843*	11	0,808*	11
30	0,524	17	0,823*	14	0,537	17	0,64*	14
Average	0.642		0.831		0.647		0.697	
Minimum	0.131		0.626		0.137		0.407	
Correct	12		20		12		23	

Table B.7: Efficiency scores for models with $\psi^x = \theta$, two choices of ψ^y, directions (G1) and (G2.0). An asterisk next to the efficiency value indicates projection onto the strongly efficient frontier. The last row indicates number of units projected onto the strongly efficient frontier.

References

Aparicio, J., Monge, J.F., 2022. The generalized range adjusted measure in data envelopment analysis: Properties, computational aspects and duality. European Journal of Operational Research 302, 621–632. doi:10.1016/j.ejor.2022.01.001.

Aparicio, J., Pastor, J.T., Ray, S.C., 2013. An overall measure of technical inefficiency at the firm and at the industry level: The 'lost profit on outlay'. European Journal of Operational Research 226, 154–162.

Aparicio, J., Pastor, J.T., Vidal, F., 2016. The directional distance function and the translation invariance property. Omega 58, 1–3.

Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science 30, 1078–1092. doi:10.1287/mnsc.30.9.1078.

Briec, W., 1997. A graph-type extension of Farrell technical efficiency measure. Journal of Productivity Analysis 8, 95–110. doi:10.1023/a:1007728515733.

Briec, W., Kerstens, K., 2009. Infeasibility and directional distance functions with application to the determinateness of the Luenberger productivity indicator. Journal of Optimization Theory and Applications 141, 55–73.

Briec, W., Lesourd, J., 1999. Metric distance function and profit: Some duality results. Journal of Optimization Theory and Applications 101, 15–33.

Chambers, R.G., Chung, Y., Färe, R., 1996a. Benefit and distance functions. Journal of Economic Theory 70, 407–419.

Chambers, R.G., Chung, Y., Färe, R., 1998. Profit, directional distance functions, and Nerlovian efficiency. Journal of Optimization Theory and Applications 98, 351–364. doi:10.1023/A:1022637501082.

Chambers, R.G., Fāure, R., Grosskopf, S., 1996b. Productivity growth in APEC countries. Pacific Economic Review 1, 181–190.

Charnes, A., Cooper, W.W., Golany, B., Seiford, L., Stutz, J., 1985. Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. Journal of Econometrics 30, 91–107. doi:10.1016/0304-4076(85)90133-2.

Chavas, J.P., Cox, T.L., 1999. A generalized distance function and the analysis of production efficiency. Southern Economic Journal 66, 294–318.

Cheng, G., Zervopoulos, P., Qian, Z., 2013. A variant of radial measure capable of dealing with negative inputs and outputs in data envelopment analysis. European Journal of Operational Research 225, 100–105.

Cooper, W.W., Park, K.S., Pastor, J.T., 1999. RAM: A range adjusted measure of inefficiency for use with additive models, and relations to other models and measures in DEA. Journal of Productivity Analysis 11, 5–42. doi:10.1023/a:1007701304281.

Cooper, W.W., Pastor, J.T., Borras, F., Aparicio, J., Pastor, D., 2011. BAM: A bounded adjusted measure of efficiency for use with bounded additive models. Journal of Productivity Analysis 35, 85–94.

Cooper, W.W., Seiford, L.M., Tone, K., 2007. Data Envelopment Analysis: A Comprehensive Text

with Models, Applications, References and DEA-Solver Software. Springer US. doi:`10.1007/978-0-387-45283-8`.

Cuesta, R.A., Zofío, J.L., 2005. Hyperbolic efficiency and parametric distance functions: with application to spanish savings banks. Journal of Productivity Analysis 24, 31–48.

Färe, R., Grosskopf, S., Lovell, C.A.K., 1985. The Measurement of Efficiency of Production. Springer Netherlands. doi:`10.1007/978-94-015-7721-2`.

Färe, R., Grosskopf, S., Margaritis, D., 2008. Efficiency and productivity: Malmquist and more, in: Fried, H.O., Knox Lovell, C.A., Schmidt, S.S. (Eds.), The Measurement of Productive Efficiency and Productivity Growth. Oxford University Press, pp. 522–622. doi:`10.1093/acprof:oso/9780195183528.003.0005`.

Färe, R., Lovell, C.A.K., 1978. Measuring the technical efficiency of production. Journal of Economic Theory 19, 150–162. doi:`10.1016/0022-0531(78)90060-1`.

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

Färe, R., Margaritis, D., Rouse, P., Roshdi, I., 2016. Estimating the hyperbolic distance function: A directional distance function approach. European Journal of Operational Research 254, 312–319. doi:`10.1016/j.ejor.2016.03.045`.

Farrell, M.J., 1957. The measurement of productive efficiency. Journal of the royal statistical society: series A (General) 120, 253–281.

Grant, M., Boyd, S., 2008. Graph implementations for nonsmooth convex programs, in: Blondel, V., Boyd, S., Kimura, H. (Eds.), Recent Advances in Learning and Control. Springer-Verlag Limited. Lecture Notes in Control and Information Sciences, pp. 95–110. `http://stanford.edu/~boyd/graph_dcp.html`.

Grant, M., Boyd, S., 2014. CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`.

Halická, M., Trnovská, M., 2019. Duality and profit efficiency for the hyperbolic measure model. European Journal of Operational Research 278, 410–421. doi:`10.1016/j.ejor.2018.12.001`.

Halická, M., Trnovská, M., 2021. A unified approach to non-radial graph models in data envelopment analysis: Common features, geometry, and duality. European Journal of Operational Research 289, 611–627. doi:`10.1016/j.ejor.2020.07.019`.

Hasannasab, M., Margaritis, D., Roshdi, I., Rouse, P., 2019. Hyperbolic efficiency measurement: A conic programming approach. European Journal of Operational Research 278, 401–409. doi:`10.1016/j.ejor.2018.12.005`.

Hudgins, L.B., Primont, D., 2007. Derivative properties of directional technology distance functions, in: Färe, R., Grosskopf, S., Primont, D. (Eds.), Aggregation, Efficiency, and Measurement. Springer, Boston, MA, pp. 31–43. doi:`10.1007/978-0-387-47677-3_2`.

Johnson, A.L., McGinnis, L.F., 2009. The hyperbolic-oriented efficiency measure as a remedy to infeasibility of super efficiency models. Journal of the Operational Research Society 60, 1511–1517.

Kerstens, K., Van de Woestyne, I., 2011. Negative data in dea: a simple proportional distance function approach. Journal of the Operational Research Society 62, 1413–1419.

Levkoff, S.B., Russell, R.R., Schworm, W., 2011. Boundary problems with the "Russell" graph measure of technical efficiency: A refinement. Journal of Productivity Analysis 37, 239–248. doi:10.1007/s11123-011-0241-3.

Mehdiloozad, M., Roshdi, I., 2014. Analyzing the concept of super-efficiency in data envelopment analysis: A directional distance function approach. doi:10.48550/arXiv.1407.2599.

Pastor, J.T., Aparicio, J., Zofío, J.L., 2022. Benchmarking Economic Efficiency: Technical and Allocative Fundamentals. International Series in Operations Research and Management Science, Springer Cham.

Pastor, J.T., Ruiz, J.L., Sirvent, I., 1999. An enhanced DEA Russell graph efficiency measure. European Journal of Operational Research 115, 596–607.

Portela, M.C.A.S., Thanassoulis, E., Simpson, G., 2004. Negative data in DEA: A directional distance approach applied to bank branches. Journal of the Operational Research Society 55, 1111–1121. doi:10.1057/palgrave.jors.2601768.

Roshdi, I., Hasannasab, M., Margaritis, D., Rouse, P., 2018. Generalised weak disposability and efficiency measurement in environmental technologies. European Journal of Operational Research 266, 1000–1012. doi:10.1016/j.ejor.2017.10.033.

Russell, R.R., Schworm, W., 2008. Axiomatic foundations of efficiency measurement on data-generated technologies. Journal of Productivity Analysis 31, 77–86. doi:10.1007/s11123-008-0119-1.

Russell, R.R., Schworm, W., 2011. Properties of inefficiency indexes on ⟨input, output⟩ space. Journal of Productivity Analysis 36, 143–156. doi:10.1007/s11123-011-0209-3.

Russell, R.R., Schworm, W., 2018. Technological inefficiency indexes: A binary taxonomy and a generic theorem. Journal of Productivity Analysis 49, 17–23. doi:10.1007/s11123-017-0518-2.

Sahoo, B.K., Mehdiloozad, M., Tone, K., 2014. Cost, revenue and profit efficiency measurement in DEA: A directional distance function approach. European Journal of Operational Research 237, 921–931. doi:10.1016/j.ejor.2014.02.017.

Ševčovič, D., Halická, M., Brunovský, P., 2001. DEA analysis for a large structured bank branch network. Central European Journal of Operations Research 9, 329–342.

Shephard, R.W., 1953. Cost and production functions princeton university press. Princeton, New Jersey .

Sueyoshi, T., Sekitani, K., 2009. An occurrence of multiple projections in DEA-based measurement of technical efficiency: Theoretical comparison among DEA models from desirable properties. European Journal of Operational Research 196, 764–794. doi:10.1016/j.ejor.2008.01.045.

Tone, K., 2001. A slacks-based measure of efficiency in data envelopment analysis. European Journal of Operational Research 130, 498–509.

Tone, K., Chang, T.S., Wu, C.H., 2020. Handling negative data in slacks-based measure data envelopment analysis models. European Journal of Operational Research 282, 926–935.