Nonconvexity in Production and Cost Functions: An Exploratory and Selective Review*

Walter Briec, Kristiaan Kerstens, and Ignace Van de Woestyne

Contents

*We acknowledge the most helpful comments of R. Chambers and G. Cesaroni on an earlier version. The usual disclaimer applies.

W. Briec
University of Perpignan, LAMPS, Perpignan, France
e-mail: briec@univ-perp.fr

K. Kerstens (✉)
IESEG School of Management, CNRS, Université de Lille, UMR 9221-LEM, Lille, France
e-mail: k.kerstens@ieseg.fr

I. Van de Woestyne
Research Unit MEES, KU Leuven, Brussel, Belgium
e-mail: ignace.vandewoestyne@kuleuven.be

S. C. Ray et al. (eds.), *Handbook of Production Economics*,
https://doi.org/10.1007/978-981-10-3450-3_15-1

Abstract

The purpose of this contribution is to provide an overview of developments in nonconvex production technologies and economic value functions, with special attention to the cost function. Apart from a somewhat selective review of theoretical issues, the emphasis is on whether the assumption of convexity makes a difference in practice. Anticipating our conclusion, we argue that traditional convex empirical results differ on average rather markedly from alternative nonconvex ones. This should make the discipline reconsider its traditional relationship with convexity in both theoretical and applied production analysis.

Keywords

Nonparametric frontier · Convexity · Production · Cost function · Scale · Productivit

Introduction

This contribution focuses on deterministic nonparametric frontier technologies that somehow relax the traditional hypothesis of convexity. Apart from developments in general equilibrium theory with nonconvexities, we are unaware of any developments in empirical production theory that allow to empirically document the eventual impact of the traditional convexity axiom. This explains the narrow and selective focus of this chapter.

The seminal article of Farrell [61] introduced a single output multiple inputs deterministic nonparametric frontier technology, but did not establish a link with linear programming. Boles [20] and Charnes et al. [39] are the first economics and operations research articles, respectively, that have given the impetus that made the nonparametric approach to production one of the great success stories in terms of both methodological developments and empirical applications. While the axiom of convexity is traditionally maintained in these nonparametric production models (see Afriat [4], Banker et al. [13], Charnes et al. [39], Diewert and Parkan [50]) as well as in the mainstream empirical economic literature on production analysis, Afriat [4] was probably the first to mention a basic single output nonconvex technology imposing the assumptions of strong input and output disposability. A multiple output version has probably been proposed for the first time in Deprins et al. [49] and these authors suggested the moniker Free Disposal Hull (FDH).

The work of Scarf [108–111] may well be considered as an important predecessor of FDH, since he studied activity analysis models based on integer data. For instance, Figure 1 displayed in Scarf [108, p. 3638] resembles the FDH as we know it. Without the pretension to recount the history of the FDH technology in detail, it suffices to mention Lovell and Vanden Eeckaut [88, footnote 2] lists another three potential historical sources of the FDH concept.

This traditional stress on convex applied production analysis is to some extent surprising, since it is theoretically well-known that important features of technology fundamentally violate the convexity of the production possibility set (see Farrell [62]). First, indivisibility implies that inputs and outputs are not necessary perfectly divisible. Furthermore, scaling down or up the entire production process in infinitesimal fractions may not be feasible. Examples include the start-up and shutdown costs in industries (see, e.g., O'Neill et al. [93] for electricity generation). Scarf [112, 113] stresses the importance of indivisibility in selecting among technological options. Second, economies of scale (e.g., modern information technology) and economies of specialization (e.g., Romer [106] on nonrival inputs in the new growth theory) violate the convexity of technology. Third, the existence of positive or negative production externalities also leads to nonconvexities. Thus, the structure of production in society is potentially full of nonconvexities.

It should be realized that the natural environment is full of nonconvexities as well (see Dasgupta and Mähler [46] for an overview). Ecologists identify pathways by which ecosystem constituents interact with one another and with the external environment. A large body of empirical work reveals that those pathways often involve transformation possibilities among environmental goods and services that constitute nonconvex sets (e.g., see Boscolo and Vincent [21] on forestry economics). In the words of Dasgupta and Mähler [46]: "The word "convexity" is ubiquitous in economics, but absent from ecology."

This book chapter is structured as follows. Section "Technologies and Distance Functions: Basic Definitions" provides some basic definitions of the traditional axioms underlying technologies and their representation via distance functions. Section "Axiom of Convexity: Arguments" discusses in detail the existing justifications for the axiom of convexity. Section "Nonparametric Nonconvex Technologies and Value Functions: Free Disposal Assumption and Minimum Extrapolation Principle" first focuses on nonconvex FDH with its extensions and the corresponding traditional convex technologies, then followed by a discussion of nonconvex economic value functions as well as efficiency decompositions and tests of convexity that have been conceived in the literature. Next, we offer an empirical perspective on the use of FDH and its extensions on a variety of topics. Finally, we discuss some further methodological refinements. Section "Mitigating Convexity: A Selection" offers a very selective review of several attempts to mitigate the impact of the convexity axiom while avoiding FDH and its extensions. Section "Conclusions" concludes and outlines some future research issues.

Technologies and Distance Functions: Basic Definitions

A production technology describes all available possibilities to transform input vectors $x = (x_1, \ldots, x_m) \in \mathbb{R}^m_+$ into output vectors $y = (y_1, \ldots, y_n) \in \mathbb{R}^n_+$. The production possibility set or technology T summarizes the set of all feasible input and output vectors: $T = \{(x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ : x \text{ can produce } y\}$. Note that it may be surprising that the main contributions in this literature continue considering that the technology is a subset of $\mathbb{R}^m \times \mathbb{R}^n$. In section "Nonparametric Nonconvex Technologies and Value Functions: Free Disposal Assumption and Minimum Extrapolation Principle" we open a perspective on considering the domain $\mathbb{N}^m \times \mathbb{N}^n$ instead.

Given our focus on input-oriented efficiency measurement later on, this technology can be represented by the input correspondence $L : \mathbb{R}^n_+ \to 2^{\mathbb{R}^m_+}$ where $L(y)$ is the set of all input vectors that yield at least the output vector y:

$$L(y) = \{x : (x, y) \in T\}. \tag{1}$$

The radial input efficiency measure is a map $E : \mathbb{R}^m_+ \times \mathbb{R}^n_+ \longrightarrow \mathbb{R}_+ \cup \{\infty\}$ that can be defined as:

$$E(x, y) = \min\{\lambda \ : \ \lambda \geq 0, \ \lambda x \in L(y)\}. \tag{2}$$

This radial efficiency measure, which is the inverse of the input distance function, indicates the minimum contraction of an input vector by a scalar λ while still remaining in the input correspondence. Obviously, the resulting input combination is located at the boundary of this input correspondence. For our purpose, the radial input efficiency has two key properties (see, e.g., Hackman [68]). First, it is smaller or equal to unity ($0 \leq E(x, y) \leq 1$), whereby efficient production on the isoquant of $L(y)$ is represented by unity and $1 - E(x, y)$ indicates the amount of inefficiency. Second, it has a cost interpretation. Note that more general efficiency measures are around in the literature: one example is the directional distance function introduced by Chambers et al. [38] that is sometimes mentioned in this contribution.

Consider a set of K observations $A = \{(x_1, y_1), \ldots, (x_K, y_K)\} \in \mathbb{R}^m_+ \times \mathbb{R}^n_+$. In the following, let us denote $\mathcal{K} = \{1, \ldots, K\}$. Nonparametric specifications of technology can then be estimated by enveloping these K observations in the set A while maintaining some basic production axioms (see Hackman [68] or Ray [104]). We are interested in defining minimum extrapolation technologies satisfying the following assumptions:

$T1$: $(0, y) \in T \Rightarrow y = 0$; $(0, 0) \in T$.
$T2$: T is closed.
$T3$: For all $(x, y) \in T$ and all $(u, v) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+$ if $(x, -y) \leq (u, -v)$, then $(u, v) \in T$.
$T4$: T exhibits $(\imath)$ constant returns to scale (CRS), $\delta T \subseteq T, \forall \delta > 0$; $(\imath\imath)$ nonincreasing returns to scale (NIRS), $\delta T \subseteq T, \forall \delta \in (0, 1)$;$(\imath\imath\imath)$ nondecreasing returns to scale (NDRS), $\delta T \subseteq T, \forall \delta > 1$; $(\imath v)$ variable returns to scale (VRS), when $(\imath)$, $(\imath\imath)$, and $(\imath\imath\imath)$ do not hold.
$T5$: T is convex.

We briefly expand on the interpretation of these basic axioms. Axiom (T1) states that there is no free lunch and that inaction is feasible. Axiom (T2) indicates that the technology is closed. Axiom (T3) represents strong or free disposability in the inputs and the outputs: inputs can be wasted without opportunity costs, and outputs can be reduced at will. Axiom (T4) defines all four traditional returns to scale hypotheses (i.e., constant, nonincreasing, nondecreasing, and variable (flexible) returns to scale). Finally, the convexity assumption (T5) is traditional, but it is not indispensable.

Axiom of Convexity: Arguments

While the axiom of convexity (T5) is traditionally maintained in economics, we develop three types of arguments to put it under scrutiny. Two arguments are related to economic theory. One argument is more pragmatic: in empirical applications, it turns out that managers often object to convexity. Sometimes the motivation to maintain the convexity axiom is just analytical convenience (see, e.g., Hackman [68, p. 2]). We think this is an argument that is valid only if one can show that convex results provide a reasonably good approximation to a potentially nonconvex economic reality.

Convexity and Duality

Often duality is invoked as a reason to maintain convexity. Since the main duality relations in economics linking, e.g., production and cost approaches presume some form of convexity, in applied empirical production analysis, researchers feel compelled to maintain the same axioms. It is an open question whether this desire for theoretical consistency is cogent.

We explore this viewpoint a little bit. The traditional duality results often fit in a general equilibrium framework that maintains convexity in its simplest forms. But, applied researchers tend to forget that general equilibrium theory has become less attractive as a general normative framework since the Sonnenschein-Mantel-Debreu results appeared in the early 1970s. Almost entirely negative conclusions appeared about the uniqueness and stability of general equilibrium. While uniqueness only occurs under restrictions void of economic realism, instability is the rule rather than the exception since almost any continuous pattern of price movements may occur in general equilibrium (see Ackerman [2]).

Furthermore, general equilibrium theory has been developed under more general conditions of nonconvexity on technology and preferences (see Chavas and Briec [41]). Realistically, this involves some process of nonlinear pricing. At the firm level, one may therefore look for proper nonconvex specifications that do justice to the nonconvexities in technology. This may imply recourse to more complex duality relations, but this is simply the price to pay for the gain in realism. The FDH and its extensions can be seen as one example that may fit into such a strategy (see, e.g., Agrell and Tind [5]).

Convexity and Time Divisibility

Several economic theorists interpret convexity of technology solely in terms of time divisibility of technologies and see no other justification for its use.

Hackman [68, p. 39] puts things clearly when discussing the axiom of convexity in his textbook:

> It does have the following "time-divisibility" justification. Suppose input vectors x_1 and x_2 each achieve output level $u > 0$. Pick a $\lambda \in [0, 1]$, and imagine operating $100\lambda\%$ of the time using x_1 and $100(1 - \lambda)\%$ of the time using x_2. At an aggregate level of detail, it is not unreasonable to assume that the weighted average input vector $\lambda x_1 + (1 - \lambda)x_2$ can also achieve output level u.

Jacobsen [70, p. 759] remarks when discussing the quasi-concavity property of the production function:

> (A.5) implies a time divisibility in the production process.

Shephard [116, p. 15] states about the property of convexity of the input set:

> Property P.8 is valid for time divisibly-operable technologies. For example, if $x \in L(u)$, $y \in L(u)$ and $\theta \in [0, 1]$, the input vector $[(1 - \theta)x + \theta y]$ may be interpreted as an operation of the technology a fraction $(1 - \theta)$ of some unit time interval with the input vector x and a fraction θ with y, assuring at least the output rate u.

The added footnote at the end of the last cited phrase reads: "Indeed the input vector $[(1 - \theta)x + \theta y]$ may have no meaning unless so interpreted."

This time divisibility argument basically ignores setup and lead times which make a switch between the underlying activities costly in terms of time. This implies that convexity becomes questionable when time indivisibilities compound all other reasons for spatial nonconvexities (e.g., indivisibilities, increasing returns to scale, economies of specialization, externalities, etc.).

Convexity and Managerial Practice: Some Skepticism Around

Decision-makers do not necessarily believe in convexity. This is evidenced in remarks, scattered in the literature, on the problems encountered in communicating the results of traditional efficiency measurement assuming convexity to decision-makers. We provide some examples of quotes reflecting this doubt of managers to the axiom of convexity.

In a study applying convex nonparametric frontier methods to measure bank branch efficiency, Parkan [96, p. 242] notes:

> The comparison of a branch which was declared relatively efficient, to a hypothetical composite branch, did not allow for convincing practical arguments as to where the inefficiencies lay.

Epstein and Henderson [53, p. 105] report similar experiences in that managers simply question the feasibility of the hypothetical projection points resulting from convex nonparametric frontiers when discussing an application to a large public-sector organization:

> The algorithm for construction of the frontier was also discussed. The frontier segment connecting A and B was considered unattainable. It was suggested that either (1) these two DMUs should be viewed as abnormal and dropped from the model, (2) certain key variables have been excluded, or (3) the assumption of linearity was inappropriate in this organization. It appears that each of these factors was present to some degree.

In a very similar vein, Bouhnik et al. [22, p. 243] state:

> Equally as important, it is our experience that managers often question the meaning of convex combinations that involve what they perceive to be irrelevant DMUs.

All quotes seem to point to the fact that convexity may well in practice combine observations that are too far apart in terms of input mix, output mix, and/or scale of operations. While one hopes for a rather uniformly dense rather well-spaced cloud of points that avoids the combination of extreme points of production, such extreme combinations apparently occur and are puzzling for managers.

In a value efficiency analysis application (a way of incorporating preference information into efficiency analysis), Halme et al. [69, p. 11] also opt for its use with FDH because this matched the preferences of management:

> The management was also more comfortable providing preference information over existing units than virtual units, and found the results valuable.

Also some researchers concede that nonconvex analysis of production facilitates the practical use of efficiency analysis. For instance, Bogetoft et al. [19, p. 859] declare in this context:

> In general, allowing the possibility set to be nonconvex facilitates the practical use of productivity analysis in benchmarking. In particular, fictitious production possibilities, generated as convex combinations of those actually observed, are usually less convincing as benchmarks, or reference units, than actually observed production possibilities.

This experience is confirmed by Halme et al. [69, p. 10]:

> During our long experience of DEA applications we repeatedly encountered the phenomenon that DMs (Decision Maker) are reluctant to evaluate other than existing units.

Obviously, we understand that this is just casual evidence that transpires from the empirical literature. But, it is useful to consider in addition to the other arguments above.

Turning to a mathematical argument, notice that there exists some general condition under which a distance function (related to the efficiency measure (2)) can characterize a nonconvex technology. This general condition is independent of the strong disposability assumption (T3) (though we use it in the remainder for computational reasons). One can provide a simple condition considering the radian subset of $R \in \mathbb{R}^d$. A subset R of $\mathbb{R}^d$ is a radian set if for all $\lambda \in [0, 1]$ and all $x \in R$, $\lambda x \in R$. Equivalently, such a subset is called a starshaped set (see Aliprantis and Border [6] for related concepts). A subset S is co-radian if for all $\lambda \geq 1$, $\lambda x \in S$. In the field of functional analysis in mathematics, a distance function is called a gauge function (analogous to the Minkowski functional for symmetrical sets). This is a function that recovers a notion of distance on a linear space. For all subset D of $\mathbb{R}^d$, the gauge function ψ_D is the map $\psi_D : \mathbb{R}^d \longrightarrow [0, \infty]$ defined by:

$$\psi_D(x) = \sup\{\delta : \delta x \in D\}, \tag{3}$$

with the convention that $\psi_D(x) = 0$ if there is no $\lambda \geq 0$ such that $\lambda x \in A$. Paralleling this definition, for all co-radian set, one can define a co-gauge as:

$$\eta_D(x) = \inf\{\delta : \delta x \in D\}. \tag{4}$$

This definition implies that for all, respectively, closed radian and co-radian sets R and S:

$$R = \{x \in \mathbb{R}^d : \psi_R(x) \geq 1\} \quad \text{and} \quad S = \{x \in \mathbb{R}^d : \eta_S(x) \leq 1\} \tag{5}$$

It follows that a production technology can be characterized from the efficiency measure (2) if and only if the input set $L(y)$ is co-radian for all $y \in \mathbb{R}^m_+$. Considering an output-oriented efficiency measure, such a characterization applies if and only if the output set is a radian (starshaped) set.

Nonparametric Nonconvex Technologies and Value Functions: Free Disposal Assumption and Minimum Extrapolation Principle

Technologies: FDH and Its Extensions

While Deprins et al. [49] are commonly acknowledged as the developers of the basic FDH model, Kerstens and Vanden Eeckaut [73] extended this basic model by introducing the possibilities of constant, nonincreasing, and nondecreasing returns to scale. This leads to the definition of three new technologies complementary to the assumption of flexible or variable returns to scale embodied in the basic FDH model.

Individual production possibility sets are based upon one production unit (x_k, y_k), the strong disposability assumption, and different maintained hypotheses of returns to scale:

$$N_\Gamma(x_k, y_k) = \left\{(x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ : x \geq \delta x_k, y \leq \delta y_k, \delta \in \Gamma\right\}, \tag{6}$$

where $\Gamma \in \{\Gamma_{CRS}, \Gamma_{NDRS}, \Gamma_{NIRS}, \Gamma_{VRS}\}$, with:

$$\begin{aligned}
&\text{(i)} \quad \Gamma_{CRS} = \{\delta : \delta \geq 0\}\,;\\
&\text{(ii)} \quad \Gamma_{NDRS} = \{\delta : \delta \geq 1\}\,;\\
&\text{(iii)} \quad \Gamma_{NIRS} = \{\delta : 0 \leq \delta \leq 1\}\,;
\end{aligned}$$

$$\text{(iv)} \quad \Gamma_{VRS} = \{\delta : \delta = 1\}.$$

Unions and convex unions of these individual production possibility sets yield the nonconvex technologies on the one hand and the traditional convex models on the other hand:

$$T_{NC,\Gamma} = \bigcup_{k \in \mathcal{K}} N_\Gamma(x_k, y_k) \text{ and } T_{C,\Gamma} = Co\Big(\bigcup_{k \in \mathcal{K}} N_\Gamma(x_k, y_k)\Big), \tag{7}$$

where Co stands for the convex hull operator.

In addition to this approach based on sets and their operations, an alternative and useful formulation can be proposed making some analogy to the traditional convex model. Let us introduce the following notation:

$$\Lambda_C = \Big\{ \sum_{k \in \mathcal{K}} z_k = 1,\ z_k \geq 0 \Big\} \text{ and } \Lambda_{NC} = \Big\{ \sum_{k \in \mathcal{K}} z_k = 1,\ z_k \in \{0, 1\} \Big\}.$$

A unified algebraic representation of convex and nonconvex technologies under different returns to scale assumptions for a sample of K observations is found in Briec et al. [30]:

$$T_{\Lambda,\Gamma} = \Big\{(x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ : (x, -y) \geq \sum_{k \in \mathcal{K}} \delta z_k (x_k, -y_k),\ z_k \in \Lambda,\ \delta \in \Gamma \Big\}, \tag{8}$$

where $\Lambda \in \{\Lambda_{NC}, \Lambda_C\}$. First, there is the activity vector (z) operating subject to a convexity (C) or nonconvexity (NC) constraint. Second, there is a scaling parameter (δ) allowing for a particular scaling of all K observations spanning the technology. This scaling parameter is smaller than or equal to 1 or larger than or equal to 1 under nonincreasing returns to scale (NIRS) and nondecreasing returns to scale (NDRS), respectively, fixed at unity under variable returns to scale (VRS), and free under constant returns to scale (CRS).

Briec et al. [30, Proposition 1] prove the following result:

Proposition 1 ([30, p. 166]). *The nonconvex technologies $T_{\Lambda_{NC},\Gamma}$ are the minimal extrapolation technologies containing the data $A = \{(x_k, y_k) : k \in \mathcal{K}\} \subset \mathbb{R}^m_+ \times \mathbb{R}^n_+$ and satisfying the axioms $T1$ to $T4$.*

The same statement for basic FDH solely has earlier been developed in Färe and Li [55]: FDH can be seen as the closest inner approximation of the true, strongly disposable but possibly nonconvex technology.

The advantages of this formulation (8) are twofold. First, it offers a coherent formulation of all basic technologies under the four basic returns to scale assumptions (T4) and under both convexity (T5) and nonconvexity. For example, under VRS (i.e., setting $\delta = 1$) and no convexity (i.e., constraint (Λ_{NC})), one obtains the classical FDH technology:

$$T_{\Lambda_{NC},\Gamma_{VRS}} = \left\{(x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ : (x, -y) \geq \sum_{k \in \mathcal{K}} z_k(x_k, -y_k), z \in \Lambda_{NC}\right\}, \quad (9)$$

as formulated by Deprins et al. [49]. As another example, under VRS and convexity (i.e., constraint (Λ_C)), one retrieves the basic technology defined by Banker et al. [13] and Färe et al. [56]:[1]

$$T_{\Lambda_C,\Gamma_{VRS}} = \left\{(x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ : (x, -y) \geq \sum_{k \in \mathcal{K}} z_k(x_k, -y_k),\ z \in \Lambda_C\right\}. \quad (10)$$

Second, its pedagogical advantage is that it neatly separates the role of the various assumptions in the formulation of technology. For instance, the restrictions on the scaling parameter (δ) relate directly to the basic definitions of the axioms on returns to scale (T4). Furthermore, the sum constraint on the activity vector z (i.e., constraint (Λ_C)) relates to the convexity axiom (T5).

In this way, one can avoid confusing statements as found in the literature. For instance, the sum constraint on the activity vector z (i.e., constraint (Λ_C)) in the envelopment or primal formulation (10) is often called a "convexity constraint" under the VRS assumption, while the CRS technology has no such constraint in the formulation of Charnes et al. [39] though it also maintains the convexity axiom (see, e.g., Cook and Seiford [44, p. 2–3]).

To compute the radial input efficiency measure (2) relative to convex technologies in (8) requires solving a nonlinear programming problem (NLP) for each evaluated observation. As shown in Briec and Kerstens [28, Lemma 2.1], this NLP can be transformed into the familiar linear programming (LP) problems that are known from the literature by substituting $w_k = \delta z_k$.

For the nonconvex technologies in (8), the radial input efficiency measure (2) requires computing a nonlinear binary mixed integer program (NLBMIP): see Briec et al. [30, p. 166]. In fact, to reduce the computational complexity of this NLBMIP problem, three distinctive alternative solution methods have been proposed in the literature. First, Podinovksi [99] reformulates all these nonconvex technologies as binary mixed integer programs (BMIP) using a big M technique. Second, starting from an existing LP model for the basic FDH model (9) (see Agrell and Tind [5]), Leleu [85] formulates for all these nonconvex technologies equivalent LP problems. Third, Briec et al. [30] develop for all nonconvex technologies an implicit enumeration strategy to obtain closed form solutions for the radial input efficiency measure (2):[2]

Proposition 2. *Let $E_{NC,\Gamma}$ denote the radial input efficiency measure defined with respect to technologies $T_{\Lambda_{NC},\Gamma}$. For all $(x, y) \in T_{\Lambda_{NC},\Gamma}$ and $k = 1, \cdots, K$, let us*

[1] Note that the convex VRS and NDRS technologies do not satisfy inaction.

[2] Note that the use of enumeration for the basic nonconvex FDH production model (9) has been around in the literature for quite a while: examples include [49, 63, 122], among others.

denote:

$$\alpha_k(x) = \max_{i \in I(x)} \frac{x_{ki}}{x_i} \quad \text{and} \quad \beta_k(y) = \max_{j \in J(y_k)} \frac{y_j}{y_{kj}},$$

where for all $(x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+$, $I(x) = \{i \in \{1, \ldots, m\} : x_i > 0\}$ *and* $J(y) = \{j \in \{1, \ldots, n\} : y_j > 0\}$. *We have, for all* $(x, y) \in T_{\Lambda_{NC},\Gamma}$*:*

$$E_{NC,\Gamma}(x, y) = \begin{cases} \min\limits_{(x_k,y_k)\in B_\Gamma(x,y)} \alpha_k(x) & \text{if} \quad \Gamma = \Gamma_{VRS}; \\ \min\limits_{(x_k,y_k)\in B_\Gamma(x,y)} \beta_k(y) \cdot \alpha_k(x) & \text{if} \ \Gamma \in \{\Gamma_{CRS}, \Gamma_{NIRS}\}; \\ \min\limits_{(x_k,y_k)\in B_\Gamma(x,y)} \{\max\{\beta_k(y), 1\} \cdot \alpha_k(x)\} & \text{if} \quad \Gamma = \Gamma_{NDRS}. \end{cases}$$

with $B_\Gamma(x, y) = \{(x_k, y_k) : \delta x_k \leq x, \delta y_k \geq y, \delta \in \Gamma\}$.

Briec and Kerstens [28, p. 148–149] refine this analysis and also offer closed form solutions for the output-oriented and graph-oriented efficiency measures. Furthermore, these authors indicate that the computational complexity of enumeration is advantageous compared to the BMIP or LP approaches. Indeed, the maximum (minimum) of a vector with n components can be calculated in the worst case in $O(n)$ arithmetic operations. Thus, to enumerate on the data set with the number of firms K, the number of arithmetic operations is about $O(LK(m + n))$, where m and n represent the number input and output dimensions and L is a measure of data storage for a given precision. A standard linear program has a $O(LK^3)$ polynomial time complexity linked to the number of observed firms K. Since $K > m + n$ in general, the time complexity of enumeration is thus better than LP. In fact, Kerstens and Van de Woestyne [75] empirically document that implicit enumeration is by far the fastest solution strategy followed by BMIP and finally LP.[3] Kerstens and Van de Woestyne [76] provide closed form solutions for the directional distance functions under alternative returns to scale assumptions.

One can mention that in this nonconvex framework, one can also treat the discrete case by considering that the technology is a subset of $\mathbb{N}^m \times \mathbb{N}^n$ (instead of $\mathbb{R}^m \times \mathbb{R}^n$). However, the radial measure (2) involves an assumption of divisibility and is therefore unsuitable. In line with Andriamasy et al. [9], one can overcome this problem by using the directional distance function (see Chambers et al. [38]) and selecting a direction that is the unit vector of $\mathbb{N}^m \times \mathbb{N}^n$.

In principle, the appropriateness of the convexity axiom can be tested for any comparison between convex and nonconvex technologies imposing a similar returns to scale hypothesis. We can define tests for the convexity of technology as a simple ratio between the convex and nonconvex input efficiency measures. Thus, the ratio:

[3] This poor performance is related to the huge size of the LP formulation in Leleu [85].

$$CT_{\Gamma}(x, y) = E_{C,\Gamma}(x, y)/E_{NC,\Gamma}(x, y) \tag{11}$$

determines a nonparametric local goodness-of-fit test for the convexity of technologies conditional on the scaling law Γ (see Briec et al. [30, p. 178]).

Economic Value Functions

The nonconvex production models have been complemented by nonconvex cost functions with corresponding specific returns to scale assumptions in Briec et al. [30]. Turning to a dual representation of technology, recall that the cost function $C : \mathbb{R}^n_+ \times \mathbb{R}^m_+ \longrightarrow \mathbb{R}_+ \cup \{\infty\}$ defines the minimum costs to produce an output vector y given a vector of semi-positive input prices ($w \in \mathbb{R}^m_+$):

$$C(y, w) = \inf\{w \cdot x \ : \ x \in L(y)\}. \tag{12}$$

Briec et al. [30, p. 175–176] establish a local duality result between the nonconvex cost functions and the nonconvex FDH and its extensions.

The computation of the cost function (12) relative to convex nonparametric technologies $T_{C,\Gamma}$ again requires an NLP to be solved for each evaluated observation. As above, this NLP can be transformed into the familiar LP problem that is known from the literature (e.g., Hackman [68]).

The cost function (12) relative to the nonconvex technology $T_{NC,\Gamma}$ involves computing a NLBMIP as mentioned above. Again, to reduce the computational complexity of this NLBMIP problem, three distinctive solution methods can be pursued. First, following the Podinovksi [99] approach, one can transform these nonconvex cost functions to BMIPs. Second, Leleu [85] formulates for all these nonconvex cost functions equivalent LP problems. Third, Briec et al. [30] develop for all nonconvex cost functions an implicit enumeration strategy yielding closed form solutions. For all $y \in \mathbb{R}^n_+$, let us denote:

$$V_{\Gamma}(y, x_k, y_k) = \left\{x \in \mathbb{R}^m_+; (x, y) \in N_{\Gamma}(x_k, y_k)\right\} \tag{13}$$

By construction, we have:

$$C_{NC,\Gamma}(y, w) = \min\left\{w \cdot x \ : \ x \in \bigcup_{k \in \mathcal{K}} V_{\Gamma}(y, x_k, y_k)\right\}. \tag{14}$$

By defining $C^{(k)}_{NC,\Gamma}(y, w) = \min\{w \cdot x : x \in V_{\Gamma}(y, x_k, y_k)\}$, we obtain:

$$C_{NC,\Gamma}(y, w) = \min_{k \in \mathcal{K}} C^{(k)}_{NC,\Gamma}(y, w). \tag{15}$$

Interestingly the above properties can be derived from the standard background of convex analysis (see Clarke [43] and Rockafeller and Wets [105] for references).[4] Given a closed subset D of $\mathbb{R}^d$, let $\delta_D : \mathbb{R}^d \longrightarrow \mathbb{R} \cup \{-\infty\}$ be the indicator function defined as:

$$\delta_D(z) = \begin{cases} 0 & \text{if } x \in D \\ -\infty & \text{if } x \notin D \end{cases} \tag{16}$$

One can then show that:

$$\inf\{w.z : z \in D\} = \inf\{w.z - \delta_D(z) : z \in \mathbb{R}^d\} = \delta_D^\star(w), \tag{17}$$

where $\delta_D^\star(w)$ stands for the conjugate of δ_D. Suppose moreover that for all $k \in \mathcal{K}$, D_k is a closed subset of $\mathbb{R}^d$ and that $D = \bigcup_{k\in\mathcal{K}} D_k$.

$$\delta_D^\star(w) = \delta^\star_{\bigcup_{k\in\mathcal{K}} D_k}(w) = \inf\{w.z - \delta_{\bigcup_{k\in\mathcal{K}} D_k}(z) : z \in \mathbb{R}^d\} \tag{18}$$

$$= \inf\{w.z - \max_{k\in\mathcal{K}} \delta_{D_k}(z) : z \in \mathbb{R}^d\} = \inf\left\{ \min_{k\in\mathcal{K}} \big(w.z - \delta_{D_k}(z)\big) : z \in \mathbb{R}^d \right\} \tag{19}$$

$$= \min_{k\in\mathcal{K}} \inf\{w.z - \delta_{D_k}(z) : z \in \mathbb{R}^d\} = \min_{k\in\mathcal{K}} \delta^\star_{D_k}(w). \tag{20}$$

Along this line we obtain for all $k \in \mathcal{K}$:

$$C^{(k)}_{NC,\Gamma}(y, w) = \delta^\star_{V_\Gamma(y,x_k,y_k)}(w) \quad \text{and} \quad C_{NC,\Gamma}(y, w) = \min_{k\in\mathcal{K}} \delta^\star_{V_\Gamma(y,x_k,y_k)}(w). \tag{21}$$

Notice that a similar method applies for efficiency analysis. The next result is then derived.

Proposition 3. *Let $C_{NC,\Gamma}(y, w)$ denote the cost function with respect to technologies $T_{\Lambda_{NC},\Gamma}$. For all $(y, w) \in \mathbb{R}^n_+ \times \mathbb{R}^m_+$, we have:*

$$C_{NC,\Gamma}(y, w) = \begin{cases} \min_{k\in\mathcal{K}} \{w \cdot x_k : y_k \geq y\} & \textit{if } \Gamma = \Gamma_{VRS}; \\ \min_{k\in\mathcal{K}} \{\beta_k(y) w \cdot x_k\} & \textit{if } \Gamma = \Gamma_{CRS}; \\ \min_{\{k:\beta_k(y)\leq 1\}} \{\beta_k(y) w \cdot x_k\} & \textit{if } \Gamma = \Gamma_{NIRS}; \\ \min_{k\in\mathcal{K}} \left\{ \max\{\beta_k(y), 1\}\, w \cdot x_k \right\} & \textit{if } \Gamma = \Gamma_{NDRS}; \end{cases}$$

[4]This point was suggested to the authors by R. Chambers.

where $J(y) = \{j : y_j > 0\}$ *and* $\beta_k(y) = \max_{j \in J(y_k)} \frac{y_j}{y_{kj}}$ *are defined as in Proposition 2.*

Remark that Ray [104, Section 10.2] shows that the basic FDH cost function yields the same result as the Weak Axiom of Cost Minimization (WACM) as defined by Varian [123]. This is intuitively obvious since WACM only imposes convexity of the input set, and thus this partial convexity yields the same cost function as the one not imposing convexity at all.

Now, there is a property of the cost function in the outputs worthwhile spelling out. Some seminal contributors to axiomatic production theory state that the cost function is nondecreasing and convex (nonconvex) in the outputs when convexity of technology is assumed (rejected) (e.g., Färe [54, p. 87], Jacobsen [70, p. 765], Shephard [116, p. 227], or Shephard [117, p. 15]). A central result established in Briec et al. [30] is that cost functions based on convex technologies are always smaller or equal to cost functions based on nonconvex technologies.

Proposition 4 ([30, p. 171]). *The convex and nonconvex cost functions* $C_{C,\Gamma}$ *and* $C_{NC,\Gamma}$*, respectively, satisfy the following properties:*

(a) *For all* $(y, w) \in \mathbb{R}^n_+ \times \mathbb{R}^m_+$, $C_{C,\Gamma}(y, w) \leq C_{NC,\Gamma}(y, w)$.
(b) *In the single output case, if* $\Gamma = \Gamma_{CRS}$*, then* $C_{C,\Gamma}(y, w) = C_{NC,\Gamma}(y, w)$.

Both cost functions are only equal in the case of CRS and a single output. Proposition 4 can be conceived as a more detailed result spelling out the precise impact of convexity on the above property of cost functions in the outputs.

Obviously, these results can also be transposed to other economic value functions. Revenue functions based upon convex technologies are higher than or equal to revenue functions based upon nonconvex technologies. Only in the single input and CRS case, both these revenue functions coincide. For the long-run profit function, by contrast, the use of convex technologies or nonconvex technologies is logically indistinguishable. However, for any other restricted profit function, one obtains the result that profit is higher or equal when tangent to a convex instead of a nonconvex technology.

Also the appropriateness of the convexity axiom can be tested by comparing convex and nonconvex value functions imposing a similar returns to scale hypothesis. A simple test of the convexity of, e.g., the cost function can be defined as a simple ratio between the convex and nonconvex cost functions. Thus, the ratio:

$$CC_{\Gamma}(y, w) = C_{C,\Gamma}(y, w)/C_{NC,\Gamma}(y, w) \tag{22}$$

determines a nonparametric local goodness-of-fit test for the convexity of cost functions conditional on the scaling law Γ (see Briec et al. [30, p. 178]). Obviously, this convexity test in Definition 22 is similar in structure to the test earlier developed in Definition 11.

Efficiency Decompositions and the Testing of Convexity: A Priori Relations

While Farrell [61] provided the first measurement scheme for the evaluation of Technical and Allocative Efficiency in a frontier context, Färe et al. [57] and Seitz [115] both offer alternative extended efficiency taxonomies. Because it is in our opinion the most widespreadly used, we stick in this contribution to the conceptual framework developed in Färe et al. [57, pp. 3–5].

The radial efficiency measure (2) used relative to different technologies entails the different concepts in this efficiency taxonomy of Färe et al. [57]. By conditioning the notation of the radial efficiency measure (2) on, e.g., a particular returns to scale hypothesis, it is straightforward to provide a formal characterization of all efficiency notions in the following definition (see, e.g., Briec et al. [30, p. 179]).

The following input-oriented efficiency notions are identified:

(a) Technical Efficiency $TE_{\Lambda}(x, y) = E_{\Lambda,\mathrm{VRS}}(x, y)$.
(b) Overall Technical Efficiency $OTE_{\Lambda}(x, y) = E_{\Lambda,\mathrm{CRS}}(x, y)$.
(c) Scale Efficiency $SCE_{\Lambda}(x, y) = E_{\Lambda,CRS}(x, y)/E_{\Lambda,VRS}(x, y)$.
(d) Overall Efficiency $OE_{\Lambda}(x, y, w) = C_{\Lambda,CRS}(y, w)/(w \cdot x)$.
(e) Allocative Efficiency $AE_{\Lambda}(x, y, w) = OE_{\Lambda}(x, y, w)/OTE_{\Lambda}(x, y)$.

While Technical Efficiency ($TE_{\Lambda}(x, y)$) requires production on the boundary of the VRS technology, Overall Technical Efficiency ($OTE_{\Lambda}(x, y)$) necessitates that production is situated on the boundary of the CRS technology. Scale Efficiency ($SCE_{\Lambda}(x, y)$) reflects a social goal and is measured by the ratio between the actual (VRS) and ideal (CRS) technological configurations. Overall Efficiency ($OE_{\Lambda}(x, y, w)$) requires computing a cost function relative to a CRS technology ($C_{\Lambda,CRS}(y, w)$) and taking the ratio between minimal and observed costs ($w \cdot x$). Allocative Efficiency ($AE_{\Lambda}(x, y, w)$) is a residual term computed by the ratio of $OTE_{\Lambda}(x, y)$ and $OTE_{\Lambda}(x, y)$.[5]

Since $E_{\Lambda,CRS}(x, y) \leq E_{\Lambda,VRS}(x, y)$, evidently $0 < SCE_{\Lambda}(x, y) \leq 1$. The embeddedness of technologies in terms of returns to scale assumptions determines the relations between these efficiency measures. These static efficiency concepts are mutually exclusive, and their radial measurement yields a multiplicative decomposition:

$$OE_{\Lambda}(x, y, w) = AE_{\Lambda}(x, y, w) \cdot OTE_{\Lambda}(x, y) \tag{23}$$

[5] This decomposition ignores structural efficiency or congestion. Recently, an attempt was made to develop new methods to measure strong forms of hypercongestion for convex and nonconvex technologies alike in Briec et al. [31]. This new methodology is empirically illustrated in Briec et al. [32]. Abad and Briec [1] transpose this methodology toward the modeling of bad outputs using a by-production framework.

where $OTE_{\Lambda}(x, y) = TE_{\Lambda}(x, y) \cdot SCE_{\Lambda}(x, y)$.

To develop tests for convexity, we clarify the relationship between convex and nonconvex decompositions:

Proposition 5 ([30, p. 180]). *For all* $(x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+$, *the relations between convex and nonconvex decomposition components are:* (a) $OTE_C(x, y) \leq OTE_{NC}(x, y)$; (b) $TE_C(x, y) \leq TE_{NC}(x, y)$; (c) $OE_C(x, y, w) \leq OE_{NC}(x, y, w)$.

Thus, while three out of the five above efficiency notions can be ordered with respect to the impact of convexity, there is no a priori ordering possible for the nonconvex and convex scale $(SCE_{\Lambda}(x, y))$ and Allocative $(AE_{\Lambda}(x, y, w))$ Efficiency components. Though the underlying efficiency measures can be ordered, it is not possible to order the ratios between these efficiency measures.

Nonparametric goodness-of-fit tests for the convexity of the efficiency components based upon constant returns to scale technologies and cost functions, respectively, are provided by the following ratios (see Briec et al. [30, p. 181]):

$$CRTE(x, y) = OTE_C(x, y)/OTE_{NC}(x, y) \tag{24}$$

and

$$CRCE_{(}x, y, w) = OE_C(x, y, w)/OE_{NC}(x, y, w). \tag{25}$$

Several methods have been proposed in the literature to obtain qualitative information regarding global returns to scale (e.g., see Seiford and Zhu [114]). Since these methods are not suitable for nonconvex technologies, Kerstens and Vanden Eeckaut [73, Proposition 2] generalize an existing goodness-of-fit method to suit all technologies. Including a fourth returns to scale case only relevant for nonconvex technologies (see Podinovksi [98]), the following proposition summarizes this method.

Proposition 6 ([35, p. 579]). *Conditional on the optimal efficient point, technology* $T_{\Lambda,VRS}$ *is globally characterized by:*

(a) *CRS* : $E_{\Lambda,NIRS}(x, y) = E_{\Lambda,NDRS}(x, y) = E_{\Lambda,VRS}(x, y)$;
(b) *IRS* : $E_{\Lambda,NIRS}(x, y) < E_{\Lambda,NDRS}(x, y) \leq E_{\Lambda,VRS}(x, y)$;
(c) *DRS* : $E_{\Lambda,NDRS}(x, y) < E_{\Lambda,NIRS}(x, y) \leq E_{\Lambda,VRS}(x, y)$;
(d) *SCRS* : $E_{\Lambda,NIRS}(x, y) = E_{\Lambda,NDRS}(x, y) < E_{\Lambda,VRS}(x, y)$;

where IRS, DRS, and SCRS stand for increasing, decreasing, and sub-constant returns to scale, respectively.

Essentially, these CRS, NIRS, and NDRS technologies are auxiliary to determine the position of an observation relative to the true flexible (i.e., VRS) returns to scale technology. Recently, Mostafaee and Soleimani-Damaneh [92] propose a

Table 1 Nonconvex and convex cost estimates: a selection

Article	Ratio $CC_{\Gamma}(y, w)$ (in %)	Remarks
Balaguer-coll et al. [11]	58.87	
Briec et al. [30]	97.76	CRS
Cummins and Zi [45]	50.55	
De Borger & Kerstens [47]	77.59	
Grifell-Tatjé & Kerstens [67]	90.85	Actual
	79.82	Ideal
Viton [124]	87.64	1 Output
	92.77	4 Outputs

more elaborated taxonomy of global returns to scale characterizations for nonconvex technologies based on results of Mostafaee and Soleimani-Damaneh [91].

Empirical Evidence on FDH and Its Extensions: The Impact of Convexity

This subsection focuses on the key question: does nonconvexity matter in empirical applications when compared to traditional convex analysis? We provide some evidence for a selection of four economic topics: (i) cost functions, (ii) efficiency decompositions, (iii) productivity growth, and (iv) capacity utilization.

Cost Function Results

In Table 1 we list a small selection of studies that report the results of convex and nonconvex frontier cost estimates. The first column lists the authors of the article, the second column reports the ratio $CC_{\Gamma}(y, w)$ as defined in Definition 22, and the third column eventually provides a remark.[6]

The Balaguer-Coll et al. [11] study on Spanish municipalities reveals that convex costs are only 58.87% of nonconvex costs at the sample average. Analyzing the US life insurance industry, Cummins and Zi [45] even report 50.55% on average for $CC_{\Gamma}(y, w)$: this means that convex cost is about half of the nonconvex costs. The De Borger and Kerstens [47] analysis of Belgian municipalities shows that convex costs are only 77.59% of convex costs. In a study of Spanish electricity distribution, Grifell-Tatjé and Kerstens [67] report a ratio of 90.85% when using data from the actual network and of 79.82% when using data from an ideal engineering network.

The Briec et al. [30] study lists a ratio of 97.76%, but this study imposes CRS and therefore meets one of the two conditions for equality (see Proposition 4).

[6] In case the study does not report cost estimates but rather overall efficiency ratios, one can obtain $CC_{\Gamma}(y, w) = C_{C,\Gamma}(y, w)/C_{NC,\Gamma}(y, w)$ by taking the ratio of the corresponding overall efficiency ratios $OE_C(x, y, w)/OE_{NC}(x, y, w)$. The observed cost in each of the denominators of $OE_{\Lambda}(x, y, w)$ cancels out.

The Viton [124] article is a bit a special case in that the author compares WACM and traditional convex cost estimates: since WACM coincides with a nonconvex estimate, this amounts to an implicit test of convexity. He reports a ratio of 87.64% under a single output specification (meeting again one of the two conditions for equality, Proposition 4) and a ratio of 92.77% under a multiple output specification.

In conclusion, it is undeniable that convexity has an important to huge impact on cost estimates and hence on Overall Efficiency.

Efficiency Decomposition

From the efficiency decomposition discussed in section "Efficiency Decompositions and the Testing of Convexity: A Priori Relations," the overall efficiency component has already been discussed in section "Cost Function Results." Therefore, we focus on technical efficiency components in this part.

As established in Proposition 5, $TE_C(x, y) \leq TE_{NC}(x, y)$. There is an abundance of studies reporting efficiency measures computed relative to basic convex (10) and nonconvex (9) technologies. We focus on just a few examples. For instance, Stroobants and Bouckaert [120] compare libraries in the Flemish region and report substantial differences between convex and nonconvex results for three specifications (though no statistical tests are reported). As another example, Mayston [90] evaluates UK economics departments and finds substantial differences at the sample level (though again no statistical tests are reported).

Cesaroni et al. [35, p. 582–583] report on the decomposition $OTE_\Lambda(x, y) = TE_\Lambda(x, y) \cdot SCE_\Lambda(x, y)$ for five secondary data sets. These authors find that convex and nonconvex $OTE_\Lambda(x, y)$ is only significantly different for two data sets, while convex and nonconvex $SCE_\Lambda(x, y)$ happens to be significantly different for all data sets and convex and nonconvex $TE_\Lambda(x, y)$ for most data sets. The same authors also focus on conflicting cases in returns to scale determination using Proposition 6: e.g., switches from increasing returns to scale (IRS) to decreasing returns to scale (DRS), from CRS to IRS, and from CRS to DRS. While one data set has no conflicting cases, four data sets find conflicting cases ranging between 6.98% and 39.02% of observations. Finally, these authors explore the markedly different patterns of ray average productivity curves under convex and nonconvex technologies.

Chavas and Kim [42, p. 69–70] report on convex and nonconvex $TE_\Lambda(x, y)$ and $SCE_\Lambda(x, y)$: while no statistical tests are reported, the descriptive statistics seem to be markedly different. Cesaroni and Giovannola [34, p. 128–129] establish results for alternative convex and nonconvex cost-based efficiency components similar to the above: though no statistical tests are mentioned, the descriptive statistics are clearly different beyond doubt.

Productivity Growth

Kerstens and Van de Woestyne [74] report empirical results for the immensely popular Malmquist productivity index (e.g., Färe et al. [60]) as well as for the Hicks-Moorsteen Total Factor Productivity (TFP) index (defined by Bjurek [17]) under various specifications of technology. For both indices, it turns out that convex and nonconvex results for both CRS and VRS yield different descriptive statistics, though no formal tests are provided regarding the statistical significance of these differences.

Kerstens and Managi [72] focus on the Luenberger productivity indicator which is defined in terms of the differences between directional distance functions (see [37]) using basic convex (10) and nonconvex (9) technologies. Analyzing a huge data set of petroleum wells, their findings can be summarized as follows. First, productivity change is on average smaller under nonconvexity, and the resulting distributions are significantly different. Second, substantially more observations tend to push the frontier outward under nonconvexity and are thus involved in creating technological change. Third, both β-convergence and σ-convergence are being tested for and happen to occur only under nonconvexity, not under the traditional convexity axiom. In a follow-up study of Chinese banks, Barros et al. [15] also find that the Luenberger productivity change is on average smaller under nonconvexity. Testing differences in productivity with respect to scale and ownership does not yield different patterns according to convexity.

Finally, Ang and Kerstens [10] study productivity of US agriculture at the state level using the Luenberger-Hicks-Moorsteen TFP indicator (introduced by Briec and Kerstens [27]) again using basic convex (10) and nonconvex (9) technologies. These authors report a higher TFP change under nonconvexity, and the resulting distributions are significantly different.

Capacity Utilization

Johansen [71] introduces the notion of plant capacity as the maximum output vector that can be produced with existing equipment with unrestricted variable inputs per unit of time. Färe et al. [59] transpose this notion into a multi-output frontier framework by using a combination of two output-oriented efficiency measures: one relative to a technology including the variable inputs and another one excluding the variable inputs. Walden and Tomberlin [125] report average output-oriented plant capacity estimates that vary between 52% and 84% in the cases of a basic convex (10) and a basic nonconvex (9) technology, respectively.

Kerstens et al. [79] argue that the output-oriented plant capacity utilization is unrealistic when the amounts of variable inputs needed to reach the maximum capacity outputs are not available. This is related to the attainability issue already noted by Johansen [71]. These authors illustrate empirically that the scaling of variable inputs is less implausible for nonconvex compared to traditional convex technologies.

Cesaroni et al. [36] define an alternative input-oriented plant capacity notion by using a combination of two sub-vector input-oriented efficiency measures only aimed at reducing the variable inputs: one relative to a standard technology and one relative to a technology with the minimum output level per dimension among all observed units. While these authors report average output-oriented plant capacity estimates that are 92% and 89% for the convex (10) and nonconvex (9) technologies, respectively, these apparent small differences nevertheless represent distributions that turn out to be statistically significantly different. For the average input-oriented plant capacity estimates, they report numbers of 120% and 121% for the convex (10) and nonconvex (9) technologies, respectively: again these apparent small differences reflect distributions that are statistically significantly different.

It goes without saying that such differences may well have potentially huge implications in the design of policies to combat overcapacity in fisheries. Kerstens et al. [77] report results from a short-run Johansen sector model allowing for the reallocation of production between firms that is developed in two steps. In the first step, output-oriented plant capacity estimates are computed. In the second step, the industry model minimizes the industry use of fixed inputs in a radial way such that total production is maintained at the current total level by reallocating production among firm capacities. From the 398 vessels in the fleet, the convex plant capacity estimates lead to maintain only 330 vessels, while the nonconvex estimates maintain 357 vessels. Thus, the required decommissioning effort resulting from the short-run Johansen sector model is larger under convexity.

Kerstens et al. [78] aim to compare empirically technical and economic capacity notions on both convex and nonconvex technologies. After defining these capacity notions, an empirical comparison is performed using a secondary data set containing data of French fruit producers. Two key empirical conclusions are that all these different capacity notions follow different distributions and also that these distributions almost always differ under convex and nonconvex technologies.

FDH and Its Extensions: Further Methodological Refinements

One can mention a whole series of methodological refinements and variations that have been introduced in the literature related to methods initially developed in a convex setting.

First, traditional radial efficiency measures in FDH models yield potentially huge amounts of slacks and surpluses since the efficient subset is limited to the corner points; nonradial input-, output-, and graph-oriented efficiency measures have been evaluated and found particularly relevant in the basic FDH model by De Borger et al. [48]. Portela et al. [101] focus on some alternative graph-oriented (or nonoriented) efficiency measures in the same context. Following up on Ebrahimnejad et al. [55] Fukuyama et al. [64] develop least-distance efficiency measures for FDH technologies that satisfy a strong monotonicity property.

Second, in the spirit of Bouhnik et al. [22] who proposed lower bound restrictions on the intensity variables to avoid unreasonable optimal activity vectors in a convex setting, Mairesse and Vanden Eeckaut [89] develop for these nonconvex production models lower and upper bound restrictions to the scaling of observations.

Third, several types of extreme points (including anchor points) can be distinguished in FDH (see Soleimani-damaneh and Mostafaee [119]). Fourth, Soleimani-damaneh [118] develops a dynamic FDH production model that can be recursively solved by means of simple enumeration.

Fifth, Tavakoli and Mostafaee [121] are the first to develop a network structure production model that opens up the black box of production via parallel and sequential production processes in a nonconvex world. These authors obtain closed form solutions for the basic efficiency measures under FDH and its extensions. Sixth, there is some work on the construction of three-dimensional sections of the efficient frontier for nonconvex models via enumeration methods as developed supra (see Krivonozhko and Lychev [80–83], Krivonozhko et al. [84]).

Finally, Tulkens [122] was the first to propose a Free Replicability Hull (FRH) by allowing for integer replications of all observations, eventually complemented by upper bounds on the integer replication process. It turns out that this FRH is computationally quite challenging (see Ehrgott and Tind [52]). In a similar vein, Green and Cook [66] define a nonconvex technology containing all observations as well as all composite observations obtained by simple aggregation. This Free Coordination Hull (FCH) can eventually also be complemented by an upper bound on the number of observations being aggregated.

Thus, most of the analysis that has been developed for convex technologies can somehow be transposed to FDH and its extensions. This simply illustrates that this rich body of analytical results is not necessarily jeopardized when opting for nonconvex technologies.

Mitigating Convexity: A Selection

It should be clear by now that if one drops the convexity axiom altogether, then FDH and its extensions are the straightforward technological and economic value function choices to consider. However, some people have sought to mitigate the impact of convexity in a variety of ways. This section offers a selection of approaches defining some alternative to the traditional convexity axiom and somehow avoiding FDH and its extensions.

Partial Convexity

Several authors have attempted to relax the convexity axiom somewhat. Petersen [97] initiated a small literature aimed at maintaining convexity in input space and in output space solely, but not in the graph of technology. The implementation of this relaxed set of assumptions is corrected by Bogetoft [18] with restrictions on the dimensionality of the production technology. Bogetoft et al. [19] relax these restrictions on the dimensionality of the input and output spaces, while Post [102] improves upon the latter article by proposing a procedure that avoids computational problems in large-scale applications.

This relaxed assumption is justified by appeal to, for instance, the law of diminishing marginal rates of substitution in the input space or to the idea of diminishing marginal rates of transformation in the output space. However, it is not clear how time divisibility can be applied in the context of this partial convexity notion. Furthermore, one may question whether there really is, for instance, a law of diminishing marginal rates of substitution in the input space. For example, Brokken [33] summarizes three studies revealing that there are increasing marginal rates of substitution of grain for roughage in beef production. Therefore, the law of diminishing marginal rates of substitution is questionable.

Podinovski [100] introduces the idea of partial convexity between certain subsets of inputs and subsets of outputs and derives BMIP for the traditional efficiency measures. Leleu [86] proposes new LP formulations combining aspects of convex and nonconvex production models across dimensions for all returns to scale assumptions and for the directional distance efficiency measure. While Podinovski [100, p. 555–556] justifies his partial convexity approach by appealing to divisibility arguments pertaining to specific inputs and/or outputs, one may wonder whether time divisibility is by definition related to the whole production process and that setup times and indivisibilities destroy convexity altogether rather than only in some subset of dimensions.

Finally, Chavas and Kim [42] adopt a different strategy to combine convex and nonconvex models by defining the technology as a union of neighborhood-based local representation of the technology each of which is convex. Obviously, the union of convex technologies needs not be convex. By choosing very small or very large neighborhoods, the technology as a union of neighborhood-based local representations of the technology converges to the nonconvex technology (9) or the convex technology (10), respectively. An obvious problem of the whole approach is the neighborhood choice and its impact on productivity and efficiency analysis.

Regular Ultra Passum Law

Olesen and Petersen [94] intend to make convex models (10) suitable to estimate optimal scale size by augmenting these with two additional maintained hypotheses which imply that the frontier is consistent with smooth curves along rays in input and in output space that obey the Regular Ultra Passum (RUP) law (i.e., monotonically decreasing scale elasticities). This RUP law implies that the production frontier must be S-shaped along any expansion path in input space. Obviously, such technologies are nonconvex in input-output space. Olesen and Petersen [94] focus on the multiple inputs single output case.

Olesen and Ruggiero [95] continue from there and focus on production technologies that are input homothetic. This allows to maintain convexity in input and in output space but to allow for nonconvexities in input-output space. This homotheticity assumption mainly serves to simplify the estimation procedure. Also this presentation assumes only one output.

In a sense, imposing the RUP law in this context again focuses on allowing for nonconvexities in input-output space, just as in section "Partial Convexity." Therefore, the same reservations prevail. Furthermore, there are long-standing misgivings on the use of homothetic structures in production theory as in Olesen and Ruggiero [95]. Already Samuelson and Swamy [107, p. 592] conclude: "Empirical experience is abundant that the Santa Claus hypothesis of homotheticity in tastes and in technical change is quite unrealistic."

From Generalized Convexity to Nonconvexity

We now focus on a modification of the $CES - CET$ model introduced by Färe et al. [58] that is a generalization of the traditional convex approach (10). This $CES - CET$ model has two parts: the output part is characterized by a *Constant Elasticity of Transformation* specification, and the input part is characterized by a *Constant Elasticity of Substitution* specification. Consider a generic map $\phi_r : \mathbb{R}^d_+ \to \mathbb{R}^d_+$ defined as $\phi_r(z) = (z_1^r, \ldots, z_d^r)$. For all $r > 0$, this function is an isomorphism from $\mathbb{R}^d_+$ to itself, and its reciprocal is defined on $\mathbb{R}^d_+$ as $\phi_r^{-1}(z) = (z_1^{1/r}, \ldots, z_d^{1/r})$. Given a subset $B = \{u_k : k \in \mathcal{K}\}_{k \in \mathcal{K}}$ of $\mathbb{R}^d_+$, from Ben-Tal [16], one can define its ϕ_r-generalized convex hull as:

$$Co^{\phi_r}(B) = \Big\{ {\phi_r}^{-1}\Big(\sum_{k \in \mathcal{K}} z_k \phi_r(u_k) \Big) : \sum_{k \in \mathcal{K}} z_k = 1, z_k \geq 0 \Big\}. \tag{26}$$

Notice that this set is not convex in the "usual" case which corresponds to the case where $r = 1$. The $CES - CET$ model can then be defined as the set:

$$\begin{aligned} T_{C,\gamma,\delta} = \Big\{ (x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ : \quad & x \geq \phi_\gamma^{-1}\Big(\sum_{k \in \mathcal{K}} z_k \phi_\gamma(x_k) \Big), \\ & y \leq \phi_\delta^{-1}\Big(\sum_{k \in \mathcal{K}} z_k \phi_\delta(y_k) \Big), \sum_{k \in \mathcal{K}} z_k = 1, z_k \geq 0 \Big\}, \end{aligned} \tag{27}$$

where γ and $\delta > 0$. Paralleling Banker et al. [13], this construction is derived from the notion of generalized convex hull defined in (26). For such a class of models, the radial efficiency measure (2) can be computed making some obvious linear transformations. Notice that Ravelojaona [103] has proposed a nonlinear version of the directional distance function (see Chambers et al. [38]) that can also be computed by linear programming methods.

Boussemart et al. [23, p. 334] state that a production technology T is said to be homogeneous of degree α if for all $\lambda > 0$:

$$(x, y) \in T \Rightarrow (\lambda x, \lambda^\alpha y) \in T. \tag{28}$$

This technology has also been termed "almost homogeneous technology of degree 1 and α." This degree of homogeneity of the technology has direct implications for the nature of returns to scale.

Proposition 7 ([23, p. 334]). *Assume that the production technology T satisfies T1–T4. Moreover, suppose that T is homogeneous of degree α.* (a) *If $\alpha > 1$, then T satisfies strictly increasing returns to scale;* (b) *if $0 < \alpha < 1$, then T satisfies strictly decreasing returns to scale.*

Thus, these homogeneous technologies exhibit either strictly increasing or strictly decreasing returns to scale according to their degree of homogeneity. Therefore, one can say that if the technology is homogeneous of degree α, then it satisfies α-returns to scale. Obviously, strictly increasing returns to scale imply nonconvexity of technology.

Boussemart et al. [23] propose to relax the definition proposed in Färe et al. [58] by considering the following production model:

$$T^{\text{alpha}}_{C,\gamma,\delta} = \Big\{(x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ : \quad x \geq \phi_\gamma^{-1}\Big(\sum_{k\in\mathcal{K}} z_k \phi_\gamma(x_k)\Big), \\ y \leq \phi_\delta^{-1}\Big(\sum_{k\in\mathcal{K}} z_k \phi_\delta(y_k)\Big), z_k \geq 0\Big\}. \tag{29}$$

where γ and $\delta > 0$. $T^{\text{alpha}}_{C,\gamma,\delta}$ satisfies an α-returns to scale assumption with $\alpha = \frac{\gamma}{\delta}$. This technology differs from the one proposed by Färe et al. [58] because it suppresses the constraint $\sum_{k\in\mathcal{K}} z_k = 1$. While their model is not compatible with an α-returns to scale assumption, model (29) satisfies axioms (T1)–(T4) and satisfies α-returns to scale under a suitable specification of α.

Proposition 8 ([23, p. 336]). *The production technology $T^{\text{alpha}}_{C,\gamma,\delta}$ defined in (27) satisfies:*

(a) *strictly increasing returns to scale if and only if $\gamma/\delta > 1$;*
(b) *strictly decreasing returns to scale if and only if $\gamma/\delta < 1$;*
(c) *constant returns to scale if and only if $\gamma/\delta = 1$;*

Furthermore, this notion of α-returns to scale has also been extended to FDH and its extensions (see Boussemart et al. [23, p. 336]).

In empirical applications, γ and δ are a priori parameters: optimal parameter values can be determined by applying a goodness-of-fit method. This can be done using a grid search method. For example, Leleu et al. [87] analyze four types of intensive care units and find overwhelming evidence of increasing returns to scale, but at the hospital level most institutions operate under decreasing returns to scale.

More recently, Boussemart et al. [24] attempt to endogenize γ and δ using global optimization tools. They propose a tractable procedure to find an optimal value of α under a generalized FDH technology. This approach fully endogenizes α and estimate its value by linear programming. For each firm $k \in \mathcal{K}$, we consider an individual technology defined by:

$$Q_{\gamma,\delta}(x_k, y_k) = \Big\{(x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ : x \geq \lambda^{1/\gamma} x_k, y \leq \lambda^{1/\delta} y_k, \lambda \geq 0\Big\}. \tag{30}$$

The global technology is then the union of individual technologies as follows:

$$T_{NC,\gamma,\delta} = \bigcup_{k\in\mathcal{K}} Q_{\gamma,\delta}(x_k, y_k). \tag{31}$$

For all $k, j \in \mathcal{K}$, let us denote:

$$E^{(k)}_{\gamma,\delta}(x_j, y_j) = \min\{\theta : (\theta x_j, y_j) \in Q_{\gamma,\delta}(x_k, y_k)\}. \tag{32}$$

By definition, one has $E^{(k)}_{\gamma,\delta}(x_k, y_k) = 1$. From Boussemart et al. [24], one can show that:

$$E^{(k)}_{\gamma,\delta}(x_j, y_j) = \Big[\beta_k(y_j)\Big]^{\delta/\gamma} . \Big[\alpha_k(x_j)\Big] \tag{33}$$

where for all k, $\alpha_k(x_j)$ and $\beta_k(x_j)$ as in Proposition 2. Notice that this result generalizes the one defined in the VRS case. It follows that:

$$E_{NC,\gamma,\delta}(x_j, y_j) = \min\{\theta : (\theta x_j, y_j) \in T_{NC,\gamma,\delta}\} \tag{34}$$

$$= \min_{k\in\mathcal{K}} \Big(\Big[\beta_k(y_j)\Big]^{\delta/\gamma} . \Big[\alpha_k(x_j)\Big]\Big). \tag{35}$$

By defining $\alpha = \gamma/\delta$, using the fact that any efficiency score is obtained in closed form, one can then find $\alpha^\star$ which maximizes the quantity M defining an index of goodness of fit as:

$$M(A;\alpha) = \prod_{k\in\mathcal{K}} E_{NC,\gamma,\delta}(x_j, y_j) = \prod_{k\in\mathcal{K}} \min_{k\in\mathcal{K}} \Big(\Big[\beta_k(y_j)\Big]^{1/\alpha} . \Big[\alpha_k(x_j)\Big]\Big) \tag{36}$$

subject to the constraint that $(x_j, y_j) \in T_{NC,\gamma,\delta}$ for all $j \in \mathcal{K}$. Taking the logarithm it is then easy to convert this optimization problem to a linear program. An empirical application is proposed in Boussemart et al. [24].

In the same vein, based on Charnes et al. [40], we now consider the piecewise Cobb-Douglas (CD) model. Let us define the map $\phi_0 : \mathbb{R}^d_{++} \longrightarrow \mathbb{R}^d_{++}$ defined as $\phi_0(u) = (\ln(u_1), \ldots, \ln(u_d))$. This function is a bijective function from $\mathbb{R}^d_{++}$ to itself, and its reciprocal is defined on $\mathbb{R}^d_{++}$ by $\phi_0^{-1}(u) = (\exp(u_1), \ldots, \exp(u_d))$. This piecewise Cobb-Douglas model can be written as:

$$T_{CD} = \Big\{(x, y) \in \mathbb{R}^{m+n}_{++} : x \geq \prod_{k\in\mathcal{K}} x_k^{\lambda_k},\ y \leq \prod_{k\in\mathcal{K}} y_k^{\lambda_\ell}, \sum_{k\in\mathcal{K}} \lambda_k = 1, \lambda \geq 0\Big\}.$$

This model is a generalized convex model derived from the notion of generalized convexity analyzed by Ben-Tal [16]. A general taxonomy is provided in the next subsection.

Semilattice Structures

In mathematics, a partially ordered set S for which every two elements have a supremum contained in S is called an upper-semilattice. Hence for some dimension $d \in \mathbb{N}$, the partial order defined by $u \leq w$ if $u_i \leq w_i$ for all $i \in \{1, \ldots, d\}$, with $u, w \in \mathbb{R}^d_+$, realizes upper-semilattice structures in $\mathbb{R}^d_+$. The supremum of u and w is determined by $u \vee w = (\max(u_1, w_1), \ldots, \max(u_d, w_d))$. Note that the operator $\vee$ can be seen as taking the component-wise maximum.

Following Briec and Horvath [25], a subset $L \subset \mathbb{R}^d_+$ is said to be a $\mathbb{B}$-convex set, if $\forall u, w \in L, \forall t \in [0, 1] : u \vee tw \in L$. Obviously, $\mathbb{B}$-convex subsets determine a special class of upper-semilattice structures in $\mathbb{R}^d_+$ of which the mathematical properties are analyzed in detail in Briec and Horvath [25]. Briec and Horvath [26] impose $\mathbb{B}$-convexity on technologies in production economics as a substitute for convexity (and nonconvexity in the sense of FDH) and study general properties of these technologies and related cost functions. Starting from the set of K observations $A = \{(x_1, y_1), \ldots, (x_K, y_K)\} \subset \mathbb{R}^m_+ \times \mathbb{R}^n_+$, the following $\mathbb{B}$-convex nonparametric technology is defined:

$$T_{\max} = \left\{ (x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ : x \geq \bigvee_{k \in \mathcal{K}} z_k x_k, \;\; y \leq \bigvee_{k \in \mathcal{K}} z_k y_k, \;\; \bigvee_{k \in \mathcal{K}} z_k = 1, \;\; z_k \geq 0 \right\}, \tag{37}$$

with the notation

$$\bigvee_{k \in \mathcal{K}} u_k = \left(\max_{k \in \mathcal{K}}(u_{k1}), \ldots, \max_{k \in \mathcal{K}}(u_{kd}) \right) \in \mathbb{R}^d_+,$$

for $u_k = (u_{k1}, \ldots, u_{kd}) \in \mathbb{R}^d_+$, $(k \in \mathcal{K})$, expanding the operator $\vee$ to multiple vectors. Notice the structural similarity with (10) by replacing summation with component-wise maximum.

Dual to the notion of an upper-semilattice, a lower-semilattice is defined as a partially ordered set S for which every two elements have an infimum contained in S. Applied to $\mathbb{R}^d_+$, this infimum of $u, w \in \mathbb{R}^d_+$ is determined by $u \wedge w = (\min(u_1, w_1), \ldots, \min(u_d, w_d))$. Obviously, the operator $\wedge$ takes the component-wise minimum of both vectors.

Using this dual notion, Adilov and Yesilce [3] define a subset $L \subset \mathbb{R}^d_+ \cup \{+\infty\}^d$ to be inverse $\mathbb{B}$-convex if $\forall u, w \in L, \forall t \in [1, +\infty] : u \wedge tw \in L$, and study its properties. By analogy with the $\mathbb{B}$-convex case, Briec and Liang [29] define the following inverse $\mathbb{B}$-convex nonparametric technology:

$$T_{\min} = \left\{ (x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ : x \geq \bigwedge_{k \in \mathcal{K}} z_k x_k, \;\; y \leq \bigwedge_{k \in \mathcal{K}} z_k y_k, \;\; \bigwedge_{k \in \mathcal{K}} z_k = 1, \;\; z_k \geq 0 \right\}, \tag{38}$$

with the notation

$$\bigwedge_{k\in\mathcal{K}} u_k = \left(\min_{k\in\mathcal{K}}(u_{k1}), \ldots, \min_{k\in\mathcal{K}}(u_{kd})\right) \in \mathbb{R}^d_+,$$

for $u_k = (u_{k1}, \ldots, u_{kd}) \in \mathbb{R}^d_+$, $(k \in \mathcal{K})$. Compared with (10), summation is now replaced with component-wise minimum. This type of production technologies allows to take into account the situation where the inputs exhibit complementarity. In such a case, the structure of the input set is similar to that of the Leontief production function.

Radial efficiency measurements can be computed with respect to both technologies $T_{\min}$ and $T_{\max}$ by using enumeration algorithms developed in Briec and Horvath [26] and Briec and Liang [29]. These new production models have recently been applied in, e.g., energy (Andriamasy et al. [7]), transportation (Barros et al. [14]), and the tourism industry (Goncalves et al. [65]).

Coming back to the model proposed by Färe et al. [58] Andriamasy et al. [8] show that these production technologies are the Painlevé-Kuratowski lower [upper] limit of the sequence of production technologies $T_{C,r,r}$ that are derived from technology $CES-CET$ (27) by setting $\gamma = \delta = r$[7]:

$$Lim_{r\longrightarrow\infty} T_{C,r,r} = T_{\max}. \tag{39}$$

In addition id $A \subset \mathbb{R}^m_{++} \times \mathbb{R}^m_{++}$

$$Lim_{r\longrightarrow-\infty} T_{C,r,r} = T_{\min}, \tag{40}$$

and finally

$$Lim_{r\longrightarrow 0} T_{C,r,r} = T_{CD}. \tag{41}$$

Andriamasy et al. [9] consider a class of closely related nonparametric production models built on the so-called Max-Plus algebra. Let us consider the semi-ring $\mathbb{R}_{\max} = \left(\mathbb{R} \cup \left\{-\infty\right\}, \oplus, \otimes\right)$ composed of the set $\mathbb{R} \cup \left\{-\infty\right\}$ which is defined by the maximization operation as addition $s \oplus t := \max(s, t)$ and the usual addition operation as multiplication $s \otimes t := s + t$. $-\infty$ and 0 are, respectively, the neutral element of the "addition" $\oplus$ and the "multiplication" $\otimes$. One can derive from this algebraic structure the following production model:

$$T_{\oplus} := \Big\{(x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ : x \geq \bigoplus_{k\in\mathcal{K}}(z_k \otimes x^k), \tag{42}$$
$$y \leq \bigoplus_{k\in\mathcal{K}}(z_k \otimes y^k), \max_{k\in\mathcal{K}} z_k = 0, z \in \mathbb{R}^K\Big\}.$$

[7]The Painlevé-Kuratowski lower [upper] limit (sometimes also called Peano limit) of the sequence of sets $\{E_n\}_{n\in\mathbb{N}}$ is denoted $Li_{n\to\infty}E_n$ $[Ls_{n\to\infty}E_n]$. For a set of points p for which there exists a sequence $\{p_n\}$ of points such that $p_n \in E_n$ for all n and $p = \lim_{n\to\infty} p_n$, a sequence $\{E_n\}_{n\in\mathbb{N}}$ of subsets of $\mathbb{R}^m$ is said to converge, in the Painlevé-Kuratowski sense, to a set E if $Ls_{n\to\infty}E_n = E = Li_{n\to\infty}E_n$, in which case we write $E = Lim_{n\to\infty}E_n$.

This model is called a Max-Plus nonparametric estimation of the production technology. The efficiency of firms can be meaningfully evaluated using the directional distance function introduced by Chambers et al. [38] for which some closed form has been provided in Andriamasy et al. [9].

Paralleling the standard technology $T_{C,CRS}$, it is quite natural to define a graph translation homothetic Max-Plus nonparametric model of the technology. This is done by dropping the last constraint in equation (42). The following technology is Max-Plus convex and satisfies a graph translation homothetic (denoted th) assumption:

$$T_{\oplus}^{\mathrm{th}} := \left\{(x, y) \in \mathbb{R}_+^m \times \mathbb{R}_+^n : \ x \geq \bigoplus_{k \in \mathcal{K}} (z_k \otimes x^k),\ y \leq \bigoplus_{k \in \mathcal{K}} (z_k \otimes y^k),\ z \in \mathbb{R}^K \right\}. \quad (43)$$

Notice that these types of algebraic structures have more recently been considered by Baldwin and Klemperer [12] to analyze discrete demand types and to prove the existence of an equilibrium with indivisibilities.

Preliminary Conclusions

This selection is by definition incomplete and somewhat subjective. For instance, we ignore Hackman [68, p. 135] who introduces the notion of projective convexity. As another example, Kleine [80] offers a series of production models with general or individual bounds on activity levels potentially leading to nonconvexities. Our limited overview just offers a perspective on a non-negligible literature seeking alternatives to the convexity axiom.

Conclusions

Section "Technologies and Distance Functions: Basic Definitions" laid the foundations by providing basic definitions of the traditional axioms underlying technologies and their representation via distance functions. Section "Axiom of Convexity: Arguments" has focused on existing justifications for the axiom of convexity. Apart from duality reasons that often seem to be misunderstood, we have stressed the time divisibility argument and its weakness when indivisibilities also affect the time dimension (e.g., setup times). Furthermore, we have cited some evidence that decision-makers often have a hard time understanding the results from convex analysis and sometimes almost explicitly object to its use.

Section "Nonparametric Nonconvex Technologies and Value Functions: Free Disposal Assumption and Minimum Extrapolation Principle" started by a discussion of the nonconvex FDH and its extensions and also their corresponding convex technologies. The focus was on computational problems related to the need to solve nonlinear binary mixed integer programs. Three solution strategies were discussed: (i) BMIP, (ii) LP, and (iii) an implicit enumeration strategy, whereby the latter turns out to be most efficient from a computational point of view. The ensuing discussion of nonconvex economic value functions also touched upon these computational problems and the same three solution strategies. Thereafter, the focus moved to some popular efficiency decomposition and the formulation of basic tests of convexity on the technology and on the cost function.

After this methodological analysis, we switched to an empirical perspective on the use of FDH and its extensions grouped under four headings: (i) cost functions, (ii) efficiency decompositions, (iii) productivity growth, and (iv) capacity utilization. A final subsection discussed a series of methodological refinements of FDH and its extensions revealing that almost all refined analysis developed for convex technologies can somehow be transposed to FDH and its extensions.

Section "Mitigating Convexity: A Selection" has offered a selective review of attempts to mitigate the impact of the convexity axiom while avoiding FDH and its extensions. We focused extensively on partial convexity, the imposition of Regular Ultra Passum laws, α-returns to scale, and semilattice structures. This review is nowhere complete and reflects our own interests and biases.

An attempt to summarize the current state of affairs may be that the alternatives for traditional convex technologies have now been around for a decade or so. Empirical results reveal that convexity matters not only for the technology but also for economic value functions. The latter may surprise some, but it reveals that the issue of imposing convexity or not cannot be taken lightly. We consider attempts to mitigate convexity while steering away from FDH and its extensions not very successful at the moment. Therefore, unless we manage to renew the axiomatic foundations of production theory in a fundamental way, it may be hard to ignore using FDH and its extensions as well as its value functions and even harder to ignore its empirical results. An open question is to what extent existing empirical methodologies need to be re-examined to be able to cope with nonconvexities: given the local nature of some of the results, new standards may need to be established. This lack of standards to report nonconvex results as well the need to go beyond traditional convex optimization that is often considered a cornerstone for economic analysis may well contribute to its negligence.

References

1. Abad A, Briec W (2019) On the axiomatic of pollution-generating technologies: non-parametric production analysis. Eur J Oper Res 277(1):377–390
2. Ackerman F (2002) Still dead after all these years: interpreting the failure of general equilibrium theory. J Econ Methodol 9(2):119–139
3. Adilov G, Yesilce I (2012) $\mathbb{B}^{-1}$-convex Sets and $\mathbb{B}^{-1}$-measurable maps. Numer Funct Anal Optim 33(2):131–141
4. Afriat S (1972) Efficiency estimation of production functions. Int Econ Rev 13(3):568–598
5. Agrell P, Tind J (2001) A dual approach to nonconvex frontier models. J Prod Anal 16(2): 129–147
6. Aliprantis C, Border K (2006) Infinite dimensional analysis: a Hitchhiker's guide, 3rd edn. Springer, Berlin

7. Andriamasy L, Barros C, Liang Q (2014) Technical efficiency of French nuclear energy plants. Appl Econ 46(18):2119–2126
8. Andriamasy L, Briec W, Mussard S (2017) On some relations between several generalized convex DEA models. Optimization 66(4):547–570
9. Andriamasy R, Briec W, Solonandrasana B (2017) Tropical production technologies. Pac J Optim 13(4):683–706
10. Ang F, Kerstens P (2017) Decomposing the Luenberger-Hicks-Moorsteen total factor productivity indicator: an application to U.S. agriculture. Eur J Oper Res 260(1):359–375
11. Balaguer-Coll M, Prior D, Tortosa-Ausina E (2007) On the determinants of local government performance: a two-stage nonparametric approach. Eur Econ Rev 51(2):425–451
12. Baldwin E, Klemperer P (2019) Understanding preferences: "Demand Types", and the existence of equilibrium with indivisibilities. Econometrica 87(3):867–932
13. Banker R, Charnes A, Cooper W (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manag Sci 30(9):1078–1092
14. Barros C, Liang Q, Peypoch N (2013) The efficiency of French regional airports: an inverse $\mathbb{B}$-convex analysis. Int J Prod Econ 141(1):668–674
15. Barros C, Fujii H, Managi S (2015) How scale and ownership are related to financial performance? A productivity analysis of the Chinese banking sector. J Econ Struct 4:article 16
16. Ben-Tal A (1977) On generalized means and generalized convex functions. J Optim Theory Appl 21(1):1–13
17. Bjurek H (1996) The Malmquist total factor productivity index. Scand J Econ 98(2):303–313
18. Bogetoft P (1996) DEA on relaxed convexity assumptions. Manag Sci 42(3):457–465
19. Bogetoft P, Tama J, Tind J (2000) Convex input and output projections of nonconvex production possibility sets. Manag Sci 46(6):858–869
20. Boles J (1966) Efficiency squared – efficient computation of efficiency indexes. In: Proceedings of the Annual Meeting (Western Farm Economics Association). Western Agricultural Economics Association, Washington, pp 137–142
21. Boscolo M, Vincent J (2003) Nonconvexities in the production of timber, biodiversity, and carbon sequestration. J Environ Econ Manag 46(2):251–268
22. Bouhnik S, Golany B, Passy U, Hackman S, Vlatsa D (2001) Lower bound restrictions on intensities in data envelopment analysis. J Prod Anal 16(3):241–261
23. Boussemart J-P, Briec W, Peypoch N, Tavéra C (2009) α-returns to scale and multi-output production technologies. Eur J Oper Res 197(1):332–339
24. Boussemart J-P, Briec W, Leleu H, Ravelojaona P (2019) On estimating optimal α-returns to scale. J Oper Res Soc 70(1):1–11
25. Briec W, Horvath C (2004) $\mathbb{B}$-convexity. Optimization 52(2):103–127
26. Briec W, Horvath C (2009) A $\mathbb{B}$-convex production model for evaluating performance of firms. J Math Anal Appl 355(1):131–144
27. Briec W, Kerstens K (2004) A Luenberger-Hicks-Moorsteen productivity indicator: its relation to the Hicks-Moorsteen productivity index and the Luenberger productivity indicator. Econ Theory 23(4):925–939
28. Briec W, Kerstens K (2006) Input, output and graph technical efficiency measures on non-convex FDH models with various scaling laws: an integrated approach based upon implicit enumeration algorithms. TOP 14(1):135–166
29. Briec W, Liang Q (2011) On some semilattice structures for production technologies. Eur J Oper Res 215(3):740–749
30. Briec W, Kerstens K, Vanden Eeckaut P (2004) Non-convex technologies and cost functions: definitions, duality and nonparametric tests of convexity. J Econ 81(2):155–192
31. Briec W, Kerstens K, Van de Woestyne I (2016) Congestion in production correspondences. J Econ 119(1):65–90
32. Briec W, Kerstens K, Van de Woestyne I (2018) Hypercongestion in production correspondences: an empirical exploration. Appl Econ 50(27): 2938–2956

33. Brokken R (1977) The case of a queer isoquant: increasing marginal rates of substitution of grain for roughage in cattle finishing. West J Agric Econ 1(1):221–224
34. Cesaroni G, Giovannola D (2015) Average-cost efficiency and optimal scale sizes in non-parametric analysis. Eur J Oper Res 242(1):121–133
35. Cesaroni G, Kerstens K, Van de Woestyne I (2017) Global and local scale characteristics in convex and nonconvex nonparametric technologies: a first empirical exploration. Eur J Oper Res 259(2):576–586
36. Cesaroni G, Kerstens K, Van de Woestyne I (2017) A new input-oriented plant capacity notion: definition and empirical comparison. Pac Econ Rev 22(4):720–739
37. Chambers R (2002) Exact nonradial input, output, and productivity measurement. Econ Theory 20(4):751–765
38. Chambers R, Chung Y, Färe R (1998) Profit, directional distance functions, and Nerlovian efficiency. J Optim Theory Appl 98(2):351–364
39. Charnes A, Cooper W, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2(6):429–444
40. Charnes A, Cooper W, Seiford L, Stutz J (1982) A multiplicative model for efficiency analysis. Socio-Econ Plan Sci 16(5):223–224
41. Chavas J-P, Briec W (2012) On economic efficiency under non-convexity. Econ Theory 50(3):671–701
42. Chavas J, Kim K (2015) Nonparametric analysis of technology and productivity under non-convexity: a neighborhood-based approach. J Prod Anal 43(1):59–74
43. Clarke F (1983) Optimization and nonsmooth analysis. Wiley, New York
44. Cook W, Seiford L (2009) Data envelopment analysis (DEA) – thirty years on. Eur J Oper Res 192(1):1–17
45. Cummins D, Zi H (1998) Comparison of frontier efficiency methods: an application to the U.S. life insurance industry. J Prod Anal 10(2):131–152
46. Dasgupta P, Mähler K-G (2003) The economics of non-convex ecosystems: introduction. Environ Resour Econ 26(4):499–525
47. De Borger B, Kerstens K (1996) Cost efficiency of Belgian local governments: a comparative analysis of FDH, DEA, and econometric approaches. Reg Sci Urban Econ 26(2):145–170
48. De Borger B, Ferrier G, Kerstens K (1998) The choice of a technical efficiency measure on the free disposal hull reference technology: a comparison using US banking data. Eur J Oper Res 105(3):427–446
49. Deprins D, Simar L, Tulkens H (1984) Measuring labor efficiency in post offices. In: Marchand M, Pestieau P, Tulkens H (eds) The performance of public enterprises: concepts and measurements, pp 243–268. North Holland, Amsterdam
50. Diewert W, Parkan C (1983) Linear programming test of regularity conditions for production functions. In: Eichhorn W, Neumann K, Shephard R (eds) Quantitative studies on production and prices, pp 131–158. Physica-Verlag, Würzburg
51. Ebrahimnejad A, Shahverdi R, Balf F, Hatefi M (2013) Finding target units in FDH model by least-distance measure model. Kybernetika 49(4):619–635
52. Ehrgott M, Tind J (2009) Column generation with free replicability in DEA. Omega 37(5):943–950
53. Epstein M, Henderson J (1989) Data envelopment analysis for managerial control and diagnosis. Decis Sci 20(1):90–119
54. Färe R (1988) Fundamentals of production theory. Springer, Berlin
55. Färe R, Li S-K (1998) Inner and outer approximations of technology: a data envelopment approach. Eur J Oper Res 105(3):622–625
56. Färe R, Grosskopf S, Lovell C (1983) The structure of technical efficiency. Scand J Econ 85(2):181–190
57. Färe R, Grosskopf S, Lovell C (1985) The measurement of efficiency of production. Kluwer, Boston

58. Färe R, Grosskopf S, Njinkeu D (1988) On piecewise reference technologies. Manag Sci 34(12): 1507–1511
59. Färe R, Grosskopf S, Valdmanis V (1989) Capacity, competition and efficiency in hospitals: a nonparametric approach. J Prod Anal 1(2):123–138
60. Färe R, Grosskopf S, Norris M, Zhang Z (1994) Productivity growth, technical progress, and efficiency change in industrialized countries. Am Econ Rev 84(1):66–83
61. Farrell M (1957) The measurement of productive efficiency. J R Stat Soc Ser A General 120(3):253–281
62. Farrell M (1959) The convexity assumption in the theory of competitive markets. J Polit Econ 67(4):377–391
63. Fried H, Lovell C, Turner J (1996) An analysis of the performance of university affiliated credit unions. Comput Oper Res 23(4):375–384
64. Fukuyama H, Hougaard J, Sekitani K, Shi J (2016) Efficiency measurement with a non-convex free disposal hull technology. J Oper Res Soc 67(1):9–19
65. Goncalves O, Liang Q, Peypoch N (2012) Technical efficiency measurement and inverse $\mathbb{B}$-convexity: Moroccan travel agencies. Tour Econ 18(3):597–606
66. Green R, Cook W (2004) A free coordination hull approach to efficiency measurement. J Oper Res Soc 55(10):1059–1063
67. Grifell-Tatjé, E., Kerstens K (2008) Incentive regulation and the role of convexity in benchmarking electricity distribution: economists versus engineers. Ann Public Cooperative Econ 79(2):227–248
68. Hackman S (2008) Production economics: integrating the microeconomic and engineering perspectives. Springer, Berlin
69. Halme M, Korhonen P, Eskelinen J (2014) Non-convex Value efficiency analysis and its application to bank branch sales evaluation. Omega 48:10–18
70. Jacobsen S (1970) Production correspondences. Econometrica 38(5):754–771
71. Johansen L (1987) Production functions and the concept of capacity. In: Førsund F (ed) Collected works of Leif Johansen, vol 1. North Holland, Amsterdam, pp 359–382
72. Kerstens K, Managi S (2012) Total Factor productivity growth and convergence in the petroleum industry: empirical analysis testing for convexity. Int J Prod Econ 139(1):196–206
73. Kerstens K, Vanden Eeckaut P (1999) Estimating returns to scale using nonparametric deterministic technologies: a new method based on goodness-of-fit. Eur J Oper Res 113(1):206–214
74. Kerstens K, Van de Woestyne I (2014a) Comparing Malmquist and Hicks-Moorsteen productivity indices: exploring the impact of unbalanced vs. balanced panel data. Eur J Oper Res 233(3):749–758
75. Kerstens K, Van de Woestyne I (2014b) Solution methods for nonconvex free disposal hull models: a review and some critical comments. Asia-Pac J Oper Res 31(1)
76. Kerstens K, Van de Woestyne I (2018) Enumeration algorithms for FDH directional distance functions under different returns to scale assumptions. Ann Oper Res 271(2):1067–1078
77. Kerstens K, Squires D, Vestergaard N (2005) Methodological reflections on the short-run Johansen industry model in relation to capacity management. Mar Res Econ 20(4):425–443
78. Kerstens K, Sadeghi J, Van de Woestyne I (2019a) Convex and nonconvex input-oriented technical and economic capacity measures: an empirical comparison. Eur J Oper Res 276(2):699–709
79. Kerstens K, Sadeghi J, Van de Woestyne I (2019b) Plant capacity and attainability: exploration and remedies. Oper Res 67(4):1135–1149
80. Kleine A (2004) A general model framework for DEA. Omega 32(1):17–23
81. Krivonozhko V, Lychev A (2017a) Algorithms for construction of efficient frontier for nonconvex models on the basis of optimization methods. Dokl Math 96(2):541–544
82. Krivonozhko V, Lychev A (2017b) Frontier visualization for nonconvex models with the use of purposeful enumeration methods. Dokl Math 96(3):650–653
83. Krivonozhko V, Lychev A (2019) Frontier visualization and estimation of returns to scale in free disposal hull models. Comput Math Math Phys 59(3):501–511

84. Krivonozhko V, Lychev A, Blokhina N (2019) Construction of three-dimensional sections of the efficient frontier for non-convex models. Doklady Math 100(2):472–475
85. Leleu H (2006) A linear programming framework for free disposal hull technologies and cost functions: primal and dual models. Eur J Oper Res 168(2):340–344
86. Leleu H (2009) Mixing DEA and FDH models together. J Oper Res Soc 60(12):1730–1737
87. Leleu H, Moises J, Valdmanis V (2012) Optimal productive size of hospital's intensive care units. Int J Prod Econ 136(2):297–305
88. Lovell C, Vanden Eeckaut P (1994) Frontier tales: DEA and FDH. In: Diewert W, Spremann K, Stehlings F (eds) Mathematical modelling in economics: essays in honor of Wolfgang Eichhorn. Springer, Berlin, pp 446–457
89. Mairesse F, Vanden Eeckaut P (2002) Museum assessment and FDH technology: towards a global approach. J Cult Econ 26(4):261–286
90. Mayston D (2014) Effectiveness analysis of quality achievements for university departments of economics. Appl Econ 46(31):3788–3797
91. Mostafaee A, Soleimani-Damaneh M (2020a) Closed form of the response function in FDH technologies: theory, computation and application. RAIRO-Oper Res 54(1):53–68
92. Mostafaee A, Soleimani-Damaneh M (2020b) Global sub-increasing and global sub-decreasing returns to scale in free disposal hull technologies: definition, characterization and calculation. Eur J Oper Res 280(1):230–241
93. O'Neill R, Sotkiewicz P, Hobbs B, Rothkopf M, Stewart W (2005) Efficient market-clearing prices in markets with nonconvexities. Eur J Oper Res 164(1):269–285
94. Olesen O, Petersen N (2013) Imposing the regular ultra passum law in DEA models. Omega 41(1):16–27
95. Olesen O, Ruggiero J (2014) Maintaining the regular ultra passum law in data envelopment analysis. Eur J Oper Res 235(3):798–809
96. Parkan C (1987) Measuring the efficiency of service operations: an application to bank branches. Eng Cost Prod Econ 12(1–4):237–242
97. Petersen N (1990) Data envelopment analysis on a relaxed set of assumptions. Manag Sci 36(3):305–314
98. Podinovski V (2004a) Local and global returns to scale in performance measurement. J Oper Res Soc 55(2):170–178
99. Podinovski V (2004b) On the linearisation of reference technologies for testing returns to scale in FDH models. Eur J Oper Res 152(3):800–802
100. Podinovski V (2005) Selective convexity in DEA models. Eur J Oper Res 161(2):552–563
101. Portela M, Borges P, Thanassoulis E (2003) Finding closest targets in non-oriented DEA models: the case of convex and non-convex technologies. J Prod Anal 19(2–3):251–269
102. Post T (2001) Estimating Non-convex production sets – imposing convex input sets and output sets in data envelopment analysis. Eur J Oper Res 131(1):132–142
103. Ravelojaona P (2019) On Constant Elasticity of Substitution – Constant Elasticity of Transformation directional distance functions. Eur J Oper Res 272(2):780–791
104. Ray S (2004) Data envelopment analysis: theory and techniques for economics and operations research. Cambridge University Press, Cambridge
105. Rockafellar R, Wets R-B (1998) Variational analysis. Springer, Berlin
106. Romer P (1990) Are nonconvexities important for understanding growth? Am Econ Rev 80(2):97–103
107. Samuelson PA, Swamy S (1974) Invariant economic index numbers and canonical duality: survey and synthesis. Am Econ Rev 64(4):566–593
108. Scarf H (1977) An observation on the structure of production sets with indivisibilities. Proc Natl Acad Sci 74(9):3637–3641
109. Scarf H (1981a) Production sets with indivisibilities Part I: generalities. Econometrica 49(1):1–32
110. Scarf H (1981b) Production sets with indivisibilities Part II: the case of two activities. Econometrica 49(2):395–423

111. Scarf H (1986a) Neighborhood systems for production sets with indivisibilities. Econometrica 54(3):507–532
112. Scarf H (1986b) Testing for optimality in the absence of convexity. In: Heller W, Starr R, Starrett S (eds) Social choice and public decision making: essays in honor of Kenneth J. Arrow, vol I. Cambridge University Press, Cambridge, pp 117–134
113. Scarf H (1994) The allocation of resources in the presence of indivisibilities. J Econ Perspect 8(4):111–128
114. Seiford M, Zhu J (1999) An investigation of returns to scale in data envelopment analysis. Omega 27(1):1–11
115. Seitz W (1971) Productive efficiency in the steam-electric generating industry. J Polit Econ 79(4):878–886
116. Shephard R (1970) Theory of cost and production functions. Princeton University Press, Princeton
117. Shephard R (1974) Indirect production functions. Verlag Anton Hain, Meisenheim am Glam
118. Soleimani-damaneh M (2013) An enumerative algorithm for solving nonconvex dynamic DEA models. Optim Lett 7(1):101–115
119. Soleimani-damaneh M, Mostafaee A (2015) Identification of the anchor points in FDH models. Eur J Oper Res 246(3):936–943
120. Stroobants J, Bouckaert G (2014) Benchmarking local public libraries using non-parametric frontier methods: a case study of Flanders. Libr Inf Sci Res 36(3–4):211–224
121. Tavakoli I, Mostafaee A (2019) Free disposal hull efficiency scores of units with network structures. Eur J Oper Res 277(3):1027–1036
122. Tulkens H (1993) On FDH Efficiency analysis: some methodological issues and applications to retail banking, courts, and urban transit. J Prod Anal 4(1–2):183–210
123. Varian H (1984) The nonparametric approach to production analysis. Econometrica 52(3):579–597
124. Viton P (2007) Cost efficiency in US air carrier operations, 1970–1984: a comparative study. Int J Transp Econ 34(3):369–401
125. Walden J, Tomberlin D (2010) Estimating fishing vessel capacity: a comparison of nonparametric frontier approaches. Mar Res Econ 25(1):23–36