



Cost functions are nonconvex in the outputs when the technology is nonconvex: convexification is not harmless

Kristiaan Kerstens¹ · Ignace Van de Woestyne²

Accepted: 1 April 2021 / Published online: 18 June 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

This contribution focuses on testing the empirical impact of the convexity assumption in estimating costs using nonparametric specifications of technology and cost functions. Apart from reviewing the scant available evidence, the empirical results based on two publicly available data sets reveal the effect of the convexity axiom on cost function estimates: cost estimates based on convex technologies turn out to be on average between 21% and 38% lower than those computed on nonconvex technologies. These differences are statistically significant when comparing kernel densities and can be illustrated using sections of the cost function estimates along some output dimension. Finally, also the characterization of returns to scale and economies of scale using production and cost functions for individual units yields conflicting results for between 19% and 31% of individual observations. The theoretical known potential impact as well as these empirical results should make us reconsider convexity in empirical production analysis: clearly, convexity is not harmless.

Keywords Data envelopment analysis · Economics · Technology · Cost analysis

JEL classification D24

We acknowledge the most helpful comments of G. Cesaroni, P. Kerstens, F. Maier-Rigaud and M. Vardanyan on earlier versions. We thank conference and seminar participants in Auckland, Berlin, Bordeaux, Edinburgh, Harbin, Helsinki, Houston, Karaj, Palmerston North, Rennes, Shanghai, Taipei and Tianjin for useful comments on earlier drafts. We are most grateful to two referees of this journal. The usual disclaimer applies. Note that this contribution has circulated under different titles.

✉ Kristiaan Kerstens
k.kerstens@ieseg.fr

Ignace Van de Woestyne
ignace.vandewoestyne@kuleuven.be

¹ IESEG School of Management, CNRS, Univ. Lille, UMR 9221-LEM, 3 rue de la Digue, F-59000 Lille, France

² Research Unit MEES, KU Leuven, Warmoesberg 26, B-1000, Brussels, Belgium

1 Introduction

The empirical analysis of production technologies or its related value functions (e.g., cost functions) are standard methods in the empirical toolbox of the applied economist. The empirical studies analysing economies of scale, elasticities of substitution, efficiency, productivity and other economic phenomena of interest serve a wide variety of academic, regulatory and managerial purposes. The traditional parametric, semi-parametric and nonparametric specifications of technologies and value functions habitually maintain the convexity axiom.

It is well-known that a variety of reasons may generate nonconvexities in technology (see Farrell (1959) for an early overview). One example is indivisibilities: the fact that inputs and outputs in production are not perfectly divisible and thus cannot be varied continuously (see Scarf 1986, 1994). Furthermore, indivisibilities limit the up- and especially the downscaling of production processes. In addition, economies of scale and economies of specialization (see, e.g., Romer (1990) on nonrival inputs in the new growth theory) as well as externalities are all well-known features violating the convexity of technology. Also the aggregation of well behaved distinct technologies (in the sense of blueprint books) may give rise to some local nonconvex range (see Hung et al. (2009)). Nonconvexities create issues about the role of prices in defining equilibria (e.g., Scarf (1986)). Forceful statements about the importance of nonconvexities in modern production technologies are found in, e.g., the work of Scarf (1994, p. 115) who maintains: “the essence of economies of scale in production is the presence of large and significant indivisibilities in production. What I have in mind are assembly lines, bridges, transportation and communication networks, giant presses and complex manufacturing plants, which are available in specific discrete sizes, and whose economic usefulness manifests itself only when the scale of operation is large.” Also Eaton and Lipsey (1997) argue that the very existence of capital goods with a lump of embodied services (rather than disembodied service flows) points to fundamental nonconvexities in production.

Nevertheless, in empirical production analysis these features have been ignored and convexity maintained because of the assumption of time divisibility (e.g., Shephard (1970)), or simply because of analytical convenience. Of course, if time is less than perfectly divisible, then also these other features can no longer be ignored and nonconvexities may well matter. Furthermore, it is -implicitly or explicitly- assumed that there is no impact of nonconvexities on the estimates of the parameters of interest in production and, e.g., cost approaches alike. Our basic contention is that the eventual impact of convexity on production and especially cost estimates has rarely if ever been explicitly tested in the empirical economic literature. Therefore, convexity can only be maintained if there is enough evidence that its impact on the majority of empirical applications is negligible. One cannot just assume that the impact of convexity on technology and cost function estimates is negligible since information on how well convex cost functions approximate nonconvex ones is missing. In the remainder, a cost function with respect to a convex/nonconvex technology is referred to as a convex/nonconvex cost function.

Therefore, the aim of this contribution is threefold. First, we summarise the limited number of studies that implicitly or explicitly tested for the impact of convexity on the cost function. Second, we empirically explore and test for the differences between estimates based on convex and nonconvex cost functions at the sample level using secondary data sources (to allow for replication), thereby focusing on the consequences for one key economic parameter of interest. In particular, the eventual differences between the characterization of economies of scale and returns to scale for convex as well as nonconvex cost functions and technologies

are illustrated. This focus on scale economies and returns to scale allows to explore whether the eventual differences between convex and nonconvex cost functions have consequences for economic parameters beyond an eventual mere scaling effect. If cost estimates and the characterization of economies of scale and returns to scale critically depend on convexity, then investment decisions regarding the upscaling or downscaling of operations may err by following the wrong signals. Third, these differences are also numerically illustrated with graphical sections and a count of change points relating convex and nonconvex cost estimates to the outputs for individual observations. To the best of our knowledge, the combination of these three aims is unique and did -to the best of our knowledge- nowhere appear in the literature.

These convex and nonconvex cost functions are estimated using nonparametric specifications for two reasons. First, there do not seem to exist alternative semi-parametric or parametric specifications that easily allow for testing convexity. Second, this nonparametric approach coincides with the nonparametric nature of the axioms under scrutiny. Or, as Fuss et al. (1978, p. 223) state: “Given the qualitative, non-parametric nature of the fundamental axioms, this suggests that the more relevant tests will be non-parametric, rather than based on parametric functional forms, even very general ones.”

Section 2 defines the technology and the cost function in general and also elaborates on the logical and empirical reasons to question the convexity assumption in both a production and cost context. Section 3 develops the empirical methodologies in detail: (i) the specific convex and nonconvex production and cost models to be estimated, (ii) how to plot sections of these cost functions and to count change points, and (iii) the methods to characterize returns to scale and economies of scale for individual observations in a convex and nonconvex setting. The method proposed to analyse economies of scale in a nonconvex (non-differentiable) setting is new in the literature. Next, Sect. 4 discusses the two data sets employed in the empirical application. Following the structure of Sect. 3, Sect. 5 presents all empirical results. Section 6 concludes.

2 Technology and cost functions: role of convexity

2.1 Technology and cost function: definitions and duality

Denoting an n -dimensional input vector $x \in \mathbb{R}_+^n$ and an m -dimensional output vector $y \in \mathbb{R}_+^m$, the production possibility set or technology T is defined as follows: $T = \{(x, y) \in \mathbb{R}_+^n \times \mathbb{R}_+^m, x \text{ can produce } y\}$. The input set associated with T denotes all input vectors x capable of producing a given output vector y : $L(y) = \{x \in \mathbb{R}_+^n : (x, y) \in T\}$.

Technology satisfies some combination of the following conventional assumptions:

(T.1) $(0, 0) \in T$ and if $(x, y) \in T$ and $x = 0 \Rightarrow y = 0$.

(T.2) T is a closed subset of $\mathbb{R}_+^n \times \mathbb{R}_+^m$.

(T.3) If $(x, y) \in T$ and $(x', y') \in \mathbb{R}_+^n \times \mathbb{R}_+^m$, then $(x', -y') \geq (x, -y) \Rightarrow (x', y') \in T$.

(T.4) $(x, y) \in T \Rightarrow \delta(x, y) \in T$ for $\delta \in \Gamma$, where:

(i) $\Gamma \equiv \Gamma^{\text{CRS}} = \{\delta : \delta \geq 0\}$;

(ii) $\Gamma \equiv \Gamma^{\text{NDRS}} = \{\delta : \delta \geq 1\}$;

(iii) $\Gamma \equiv \Gamma^{\text{NIRS}} = \{\delta : 0 \leq \delta \leq 1\}$;

(iv) $\Gamma \equiv \Gamma^{\text{VRS}} = \{\delta : \delta = 1\}$.

(T.5) T is convex, i.e., for all $(x_1, y_1) \in T$, $(x_2, y_2) \in T$, $\lambda \in [0, 1]$: $\lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2) \in T$.

Traditional axioms on technology include: (i) inaction is feasible, and there is no free lunch, (ii) closedness, (iii) free disposal of inputs and outputs, (iv) returns to scale assumptions (i.e., constant returns to scale (CRS), nondecreasing returns to scale (NDRS), nonincreasing returns to scale (NIRS), and variable returns to scale (VRS)), and (v) convexity of technology (see, e.g., Färe (1988), Hackman (2008) or Ray (2004) for details). Key assumptions distinguishing some of the technologies in the empirical analysis are convexity and CRS.

While axioms (T.1) – (T.2) are considered regularity conditions, the other axioms are debatable. Free disposal or monotonicity precludes phenomena like congestion to become visible. Monotonicity violations are known to affect duality relations (see Barnett 2002) and occur rather often (see, e.g., Sauer 2006 for a review). A sharp criticism of the CRS hypothesis is found in Scarf (1994). In this contribution, CRS as well as NIRS and NDRS technologies are auxiliary to measuring returns to scale and economies of scale (details are discussed in Sect. 3.3). Finally, several reasons to doubt convexity are discussed in Sects. 2.2 and 2.3.

Note that every technology T^{NC} not satisfying the convexity axiom (T.5) can be extended to a convex technology T^C by taking the convex hull. This convexified technology is defined by $T^C = \{\lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2) \mid \forall(x_1, y_1) \in T^{NC}, \forall(x_2, y_2) \in T^{NC}, \forall\lambda \in [0, 1]\}$.

The input distance function completely characterizes the input set $L(y)$ and is defined as follows:

$$D_i(x, y \mid T) = \max\{\lambda : \lambda \geq 0, (x/\lambda, y) \in T\} = \max\{\lambda : \lambda \geq 0, x/\lambda \in L(y)\}. \tag{1}$$

Its main properties are: (i) $D_i(x, y \mid T) \geq 1$, with efficient production on the boundary (isoquant) of $L(y)$ represented by unity; (ii) it has a cost interpretation (see, e.g., Färe (1988) or Hackman (2008)). The inverse of the input distance function $E_i(x, y \mid T) = [D_i(x, y \mid T)]^{-1}$ determines the radial input efficiency measure. If $\delta^* = E_i(x, y \mid T)$ for some observation $(x, y) \in T$, then $(\delta^*x, y) \in T$ is referred to as the optimal projected point in the input direction.

Turning to a dual representation of technology, the cost function is defined as the minimum expenditures to produce an output vector y given a vector of semi-positive input prices ($w \in \mathbb{R}_+^n$):

$$C(y, w \mid T) = \min\{wx : (x, y) \in T\} = \min\{wx : x \in L(y)\}. \tag{2}$$

Duality relations link primal and dual formulations of technology. Duality allows a well-behaved technology to be reconstructed from the observations on cost minimizing producer behavior, and the reverse. Following (Hackman 2008, p. 114) the duality between input distance function (1) and cost function (2) is:

$$D_i(x, y \mid T) = \min_w \{wx : C(y, w \mid T) \geq 1\}, x \in L(y) \tag{3}$$

$$C(y, w \mid T) = \min_x \{wx : D_i(x, y \mid T) \geq 1\}, w > 0. \tag{4}$$

Traditionally, such duality relation is established under the hypothesis of a convex input set solely (see Färe (1988, p. 84), Hackman (2008, pp. 114–116: Theorem 7.15), or Jacobsen (1970, pp. 760, 768) for details): this is clearly a weaker version of (T.5). Thus, traditional cost functions estimated on a convex (T.5) technology then clearly impose more structure than strictly required by duality.

The Weak Axiom of Cost Minimisation (WACM) is a well-known consistency test for cost minimization that only imposes strong disposal of inputs and outputs (Varian 1984, pp. 580–582). Brieu et al. (2004) are among the first to establish a local (instead of global) duality result between nonconvex technologies that only impose strong disposal and obey

different scaling laws and the corresponding nonconvex cost functions. Ray (2004) shows that the basic nonconvex cost function with VRS is equivalent to WACM. Thus, traditional cost functions estimated on a convex ($T.5$) technology impose more structure and are not compatible with WACM.

Therefore, it remains an open question whether the traditional cost functions estimated on convex ($T.5$) technologies are different from these nonconvex cost functions.

2.2 Logical arguments questioning convexity

A logical critique of the convexity assumption can consider a variety of arguments.

First, there are some properties of the cost function in the outputs worthwhile spelling out. Some seminal contributions to axiomatic production theory indicate that the cost function is convex in the outputs if and only if technology is convex (e.g., Jacobsen (1970, Corollary 5.5)). Consequently, using the logic rule of contraposition, the cost function is nonconvex if and only if technology is nonconvex. Note that nonconvex here must be interpreted as “not convex”.¹ This property of the cost function is ignored in most advanced micro-economic textbooks (see Mas-Colell et al. (1995, p. 141), but see Kumbhakar and Lovell (2000, p. 34) for an exception). This empirical property has to our knowledge never been explicitly tested and has been sharpened by Brieu et al. (2004) who establish that (i) cost functions estimated on convex technologies yield lower or equal cost estimates compared to cost functions estimated on nonconvex technologies; and (ii) both types of cost functions are only identical for the single output and CRS case:

Proposition 2.1 (Brieu, Kerstens, and Vanden Eeckaut, 2004, p. 171) *Cost functions with respect to a nonconvex technology T^{NC} and its convexified technology T^C satisfy the following properties:*

- (a) *in general: $C(y, w | T^C) \leq C(y, w | T^{NC})$ for a given Γ ;*
- (b) *in the case of $\Gamma = \Gamma^{CRS}$ and a single output: $C(y, w | T^C) = C(y, w | T^{NC})$.*

These properties of the cost function have been ignored in most if not all empirical studies.

Second, the motivation to maintain the convexity axiom is either analytical convenience (e.g., Hackman 2008, p. 2), or time divisibility. Indeed, Hackman (2008, p. 39), Jacobsen (1970, p. 759) and Shephard (1970, p. 15) interpret convexity of technology solely in terms of time divisibility of technologies and see no other justification for its use. This time divisibility argument ignores setup and lead times that make switching between the underlying activities costly. The empirical evidence on the relevance of setup times in industrial scheduling activities is vast (see Allahverdi et al. (2008) for a recent survey). Recent studies on the time allocation between different, progressively assigned projects reveal that task juggling in services is widespread despite its induced productivity losses (e.g., Coviello et al. (2014)). While call setup times seem negligible today, minor alleged differences in communication speeds between participants are at the heart of academic and policy discussions on the role of high frequency trading in today’s financial markets (e.g., Savani (2012)).² Thus, perfect time divisibility is a questionable assumption that ideally needs empirical validation. While all reasons commonly advanced to explain nonconvexities in production are features of the

¹ The same reasoning applies to the revenue function and all variations of the profit function, except of course the long run profit function where the effect of convexity or not cannot logically be distinguished.

² Budish et al. (2014) claim that the current defects in market microstructure are linked to a continuous-time market design based on a continuous limit order book. Instead, they propose a discrete-time frequent batch auction.

real economy, these can be ignored when modeling production and costs if and only if time is perfectly divisible. If the time dimension is imperfectly divisible, then these other reasons for nonconvexities exacerbate this time indivisibility and it is an open question to which extent nonconvexities can be ignored in the empirical modeling of production and costs.

Third, note that the exact nature of the duality results does not make a difference. Even if duality between distance and cost function is established under convexity of the input set only as in Varian (1984) (rather than a convex technology) and one employs corresponding empirical methodologies only imposing convexity of the input sets, then the above arguments remain valid: if convexity of technology is implausible, then this also applies to any form of partial convexity. In particular, time divisibility is inconceivable as a legitimization of partial convexity since it applies to the underlying production activities: time divisibility of the input vectors independent of the resulting output vectors (or the reverse) makes little sense. By Occam's razor, we adopt the simplest theory and question convexity entirely.³ Obviously, a nonconvex or a convex input set yield the same cost function (see Varian (1984)),⁴ but a nonconvex or a convex output set (or technology) would not yield the same cost function, as indicated by the above properties of the cost function.

Finally, sometimes convexity is not considered a primitive axiom, but it is implied by additivity and divisibility. But, additivity and divisibility do not only imply convexity, but also CRS. This CRS assumption is at odds with indivisibilities and the lower bounds on the scaling of almost all production processes (see *supra* and Scarf (1994) for a sharp critique).⁵

2.3 Empirical arguments questioning convexity

While nonconvexities are often mentioned in economic theories, evidence on its incidence on production and cost estimates is relatively rare. We here draw on two different empirical literatures providing some evidence as to its existence. Among the exceptions, one can either mention an abundant amount of studies from the specialised engineering literature, or a very limited amount of studies published in the economic literature.

³ Bogetoft et al. (2000) is the first article maintaining convexity on input and output space, but relaxing convexity in the graph of technology without restrictions on the dimensionality of input and output spaces. Li (2019) adds variable returns to scale to this quasiconcavity of production, but only manages to treat the single output case. This framework is justified by appeal to, for instance, the law of diminishing marginal rates of substitution in inputs space, or to the idea of diminishing marginal rates of transformation in outputs space. But, one may question whether there really is a law of diminishing marginal rates of substitution in inputs space. For example, Brokken (1977) summarises three studies revealing that there are increasing marginal rates of substitution of grain for roughage in beef production: thus, establishing nonconvexities in input space.

⁴ Just as a nonconvex or a convex technology yields the same profit function. This implies that convexity cannot logically be tested using the profit function (even though any variation on the long-run profit function would allow testing for convexity).

⁵ If one is willing to accept CRS, then the plausibility of additivity and divisibility separately is at stake. First, perfect divisibility of inputs and/or outputs is probably the most debatable assumption. Many if not most operations management problems in industry and distribution involve some form of indivisibilities and input fixities resulting in complex integer and possibly nonlinear optimization problems. In general, all production processes seem to have some lower limit below which a process cannot possibly be scaled down realistically. Therefore, divisibility is highly questionable (see Scarf (1994) or Winter (2008) for more detailed criticisms). Second, while additivity is essential to define free entry and may seem plausible at first, it is not beyond criticism since it presupposes spatial separation and noninteraction (see Winter (2008)). Since additivity relates to the aggregation of results of activities occurring in geographically distinct places, transportation costs must be small in order to be safely ignored. When activities are close for transportation costs to be negligible, then the risk of production externalities looms when activities get "too close" to create interactions. Furthermore, location is important in the definition of quite some outputs (e.g., Italian and Californian lemons are considered different).

First, the empirical evidence on engineering cost estimates available on some sectors (e.g., electricity generation) points to a nonnegligible impact of convexity in principle. For instance, nonconvexities in electricity generation due to minimum up and down time constraints, multi-fuel effects, etc. leading to nonconvex and nondifferentiable variable costs are amply documented in, e.g., Bjørndal and Jörnsten (2008). However, comparisons with convex estimates are unknown to us. Seemingly, nonconvex models are a necessity and such comparisons have no use for engineers.

Second, in the economic literature one can distinguish between production and cost studies investigating convexity. Though less relevant for our purpose, there is slightly more evidence on the impact of the convexity axiom on technology-related estimates. Therefore, we briefly summarise some evidence drawn from several streams in the production literature.

First, Kerstens and Managi (2012) analyze technical change in a large panel of oil field petroleum data using a Luenberger productivity indicator: they report substantial productivity differences depending on whether convex or nonconvex technologies are used in the estimation and find only under nonconvexity that initial poor productivity fields experience faster growth than initial high productivity fields (β -convergence) and a reduction in the dispersion of productivity levels (σ -convergence). More or less in a similar vein (Alam and Sickles 2000), for instance, examine time series data on the convergence of nonconvex technical efficiency in the US airline industry.

Second, another literature offering indications on nonconvexities are parametric production studies using flexible functional forms and testing for the satisfaction of monotonicity and/or curvature conditions. If curvature or monotonicity conditions are violated, then standard second-order conditions for optimizing behavior fail to hold and duality relations break down (e.g., Barnett 2002). Sauer (2006) revisits eight parametric frontier studies in agriculture. Apart from monotonicity violations, he finds evidence of violations of the law of diminishing marginal productivity for one to four of the maximally eight considered inputs. Furthermore, none of the estimated production frontiers fulfills the curvature criterion of quasi-concavity. One interpretation of these curvature failures is that these reveal the existence of nonconvexities.

Furthermore, it is widely acknowledged that many operations management problems in industry and distribution involve some form of indivisibilities requiring integer optimization. These technology-related findings are not surprising and are reminiscent of arguments made by engineering production function advocates stating that most production processes only under strict conditions result in well-behaved neo-classical technologies (see Wibe 1984).

Turning now to cost studies, the evidence is rather scarce. Apart from some scant evidence on the relevance of nonconvexities in cost function estimates (e.g., Hasenkamp 1976 reporting economies of scale, and also economies of specialization for flexible enough functional forms, or Izadi et al. (2002) who report estimates compatible with a nonconvex iso-cost output possibility set), to our knowledge few studies have documented any differences between convex and nonconvex cost estimates.⁶ To the best of our knowledge, a new and complete list of cost studies testing convexity is provided in Table 1. While the first column contains the author(s), the second column lists the percentage difference between nonconvex and convex costs relative to nonconvex costs (i.e., $(C(y, w | T^{NC}) - C(y, w | T^C))/C(y, w | T^{NC})$), while the third column adds some qualifying remarks.⁷ All studies have multiple inputs and outputs, unless otherwise indicated in the third column. Note that only few studies mention

⁶ The Hasenkamp (1976) results are at least partially attributable to the strong assumptions on input/output separability, among others.

⁷ While some studies offer explicit comparisons, some results in Table 1 are based on own computations. In particular, some studies report cost efficiency ratios based on cost frontier estimates under convexity and

explicitly that there is an issue of convexity at stake distinguishing these results. We briefly list key results per study in alphabetic order.

We start with the studies contrasting convex and nonconvex nonparametric frontier models (see Sect. 3.1). Balaguer-Coll et al. (2007) analyse Spanish local governments and report a 41.14% difference at the sample level. This study also reports this percentage for four classes in terms of population size: it monotonously decreases from 49.59% for small municipalities (population below 10,000) to 13.38% (population above 20,000). Bricet et al. (2004) study the broad-acre Western Australian agricultural sector over the period 1953 to 1987 and obtain a mild 2.24% difference at the sample level under CRS (i.e., under one of the conditions that make both results coincide). Cummins and Zi (1998) focus on 445 US life insurance companies over a balanced panel covering the period 1988–1992 and reveal a huge 49.45% difference at the sample level. These differences are also reported per year and vary between 45.65% in 1992 to 54.35% in 1988. De Borger and Kerstens (1996) report a 22.41% difference for all 589 Belgian local governments in 1985. Garbaccio et al. (1994) report results for the US Savings and Loan industry in 1986: contrasting WACM and traditional convex cost frontiers, they find a difference of 23.46% and 22.37% for firms that remain solvent and insolvent, respectively, in the years 1988 to 1990.⁸ A probit analysis reveals that the nonconvex cost efficiency results yield the most accurate prediction of insolvency for both the least and the most most efficient firms. Grifell-Tatjé and Kerstens (2008) analyse both observed data and ideal engineering data for Spanish electricity distributors in 1996 and find the largest difference for the latter data, probably because engineers model an ideal grid implicitly or explicitly using nonconvex models. Narbón-Perpiñá et al. (2019) also report on Spanish local governments using three models: one with compulsory minimal services, one with additional services on top of minimal services, and one with all services weighted by a quality indicator. In 2008, this leads to a 33.60%, 42.42% and 25.44% difference for minimal services, all services, and all services discounted with quality respectively. The same authors also report these results for the years 2009–2013, but there is not much variation over time and therefore we do not report these results. Narbón-Perpiñá et al. (2020) analyse Spanish local governments of municipalities between 1000 and 50,000 inhabitants for the 2008–2015 period: this period leads to a 28.29% difference. Ray and Mukherjee (1995) report results for 123 US electric utilities and finds a difference of 11.02% in this single output case. Resti (1997) presents results for a panel of 270 Italian banks over the period 1988–1992 and finds a difference varying between a minimum of 20.73% and a maximum of 22.98% over these years.⁹ Viton (2007) analyses US air carriers over the period 1970–1984 and lists a difference at the sample level of 7.29% in the single output case (i.e., under one of the conditions that make both results coincide) and 12.45% in the case of four outputs when contrasting WACM and traditional convex cost frontiers.

We end with three peculiar cases. First, Ang et al. (2018) compute dynamic variable profit frontiers without and with land reallocation for a balanced panel of 44 English and Welsh farms over the seven years 2007–2013 and find an average difference of 19.12% and 28.55% respectively over the years.¹⁰ Second, car manufacturing costs are known to be nonconvex

Footnote 7 continued

nonconvexity. Taking a ratio of these efficiency ratios nets out the observed cost and reveals the difference in cost estimates under convexity and nonconvexity.

⁸ These authors also report results for six size classes each time for solvent and insolvent firms separately: the difference varies between a minimum of 9.38% and a maximum of 29.63%.

⁹ Nonconvex results are presented graphically. Our numbers are based on information provided by the author.

¹⁰ For the dynamic profit frontiers we report the difference between convex and nonconvex results relative to nonconvex results. We thank F. Ang for providing these results in personal communication.

due to changes in the number of shifts and in the temporary shut down of plants. Copeland and Hall (2011, p. 246) develop a specific dynamic engineering cost model of motor vehicle production in a single plant allowing for changes in price, labour and inventory stocks and report a modest 4.36% higher average nonconvex cost per vehicle compared to a convex alternative specification. Third, Ray (1997) contrasts a nonconvex nonparametric frontier cost estimate with the one resulting from a convex stochastic parametric cost frontier and reports a 24.73% difference.

One obvious conclusion from Table 1 is the percentage difference between the convex and nonconvex costs relative to nonconvex costs varies between a mild 2.24% under CRS to a huge 49.45% at the sample level. When ignoring studies applying one of the conditions that make both results coincide, this range varies between about 10% to about 54% in one particular case. The order of magnitude of these differences is large enough to justify digging deeper into this matter.

Since the vast majority of empirical approaches in production and cost analysis implicitly or explicitly maintains convexity of technology, one may conjecture that academics consider the convex approach a good approximation to a nonconvex approach. But, the quality of this approximation remains under-explored. If the studies in Table 1 are confirmed, then the true technology is nonconvex and cost function estimates on convex technologies are seriously downwardly biased.

Most studies listed in Table 1 focus on reporting results based on several methodologies. None of these studies verifies whether these different cost results lead to different conclusions with regard to key economic concepts and parameters of interest. To the best of our knowledge, our study is the first to move beyond the simple methodological contrasting of cost estimates by analysing the economic consequences in terms of returns to scale and economies of scale.

2.4 Conclusion: questioning convexity

Wrapping up, duality relations do not need a convex ($T.5$) technology. Cost functions differ under convexity and nonconvexity. Furthermore, convexity is not a plausible assumption since perfect time divisibility is questionable. Finally, the scarce empirical evidence on cost functions indicates potentially huge impacts of convexity. This needs further exploration.

3 Empirical methodology

3.1 Nonparametric specifications of technology and cost functions

In the nonparametric approach to production theory convexity of technology is traditionally maintained (see Afriat 1972 or Diewert and Parkan 1983). Here, we focus on the development of a nonconvex series of technologies and cost functions. Starting from a free disposal hull imposing free disposal of inputs and outputs but no convexity (see Afriat 1972 for the single output case), a first extension is the introduction of specific returns to scale assumptions into this basic model and the proposal of a new goodness-of-fit method to infer the characterization of returns to scale for nonconvex technologies (see Kerstens and Vanden Eeckaut 1999). Another step extending the scope for nonconvex production modeling is the proposal of nonconvex cost functions with specific returns to scale assumptions (Briec et al. 2004). Some additional nonparametric, nonconvex production models are discussed in, e.g., Hackman

Table 1 Nonconvex and convex cost estimates: literature review

Δ w.r.t. NC (%)	Remarks
<i>Nonparametric frontier studies:</i>	
Balaguer-Coll et al. (2007)	
41.13	Sample
49.59	pop [†] < 1000
46.14	1000 ≤ pop < 5000
27.35	5000 ≤ pop < 20000
13.38	pop ≥ 20000
Briec et al. (2004)	
2.24	CRS
Cummins and Zi (1998)	
49.45	Sample
54.35	1988
49.45	1989
50.00	1990
47.83	1991
45.65	1992
De Borger and Kerstens (1996)	
22.41	
Garbaccio et al. (1994)	
23.46	Solvent
22.37	Insolvent
Grifell-Tatjé and Kerstens (2008)	
9.15	Actual
20.18	Ideal
Narbón-Perpiñá et al. (2019)	
33.60	Minimal services
42.42	All services
25.44	All services and quality
Narbón-Perpiñá et al. (2020)	
28.19	Average 2008–2015
Ray and Mukherjee (1995)	
11.02	1 Output
Resti (1997)	
21.71	1988
20.73	1989
22.70	1990
22.12	1991
22.98	1992
Viton (2007)	
7.29	1 Output
12.45	4 Outputs

Table 1 continued

Δ w.r.t. NC (%)	Remarks
<i>Alternative studies:</i>	
Ang et al. (2018)	
19.12	Dynamic profit max. without land reallocation
28.55	Dynamic profit max. with land reallocation
Copeland and Hall (2011)	
4.36	Dynamic motor vehicle production model
Ray (1997)	
24.73	1 Output

[†]pop = municipality population

(2008). The remainder of this subsection draws on Briec et al. (2004), unless otherwise indicated.

A unified algebraic representation of convex and nonconvex technologies under different returns to scale assumptions is possible as follows:

$$T^{\Lambda, \Gamma} = \left\{ (x, y) : x \geq \sum_{k=1}^K x_k \delta z_k, y \leq \sum_{k=1}^K y_k \delta z_k, z \in \Lambda, \delta \in \Gamma \right\}, \tag{5}$$

where

(i) $\Gamma \equiv \Gamma^{\text{CRS}} = \{\delta : \delta \geq 0\}$; (ii) $\Gamma \equiv \Gamma^{\text{NDRS}} = \{\delta : \delta \geq 1\}$; (iii) $\Gamma \equiv \Gamma^{\text{NIRS}} = \{\delta : 0 \leq \delta \leq 1\}$; (iv) $\Gamma \equiv \Gamma^{\text{VRS}} = \{\delta : \delta = 1\}$;

and

(i) $\Lambda \equiv \Lambda^{\text{C}} = \left\{ z : \sum_{k=1}^K z_k = 1 \text{ and } z_k \geq 0 \right\}$;

(ii) $\Lambda \equiv \Lambda^{\text{NC}} = \left\{ z : \sum_{k=1}^K z_k = 1 \text{ and } z_k \in \{0, 1\} \right\}$.

There is one activity vector z operating subject to a nonconvexity or convexity constraint and a scaling parameter δ allowing for a particular scaling of all K observations determining the technology. The activity vector z of real numbers summing to unity represents the convexity axiom, while the same sum constraint with each vector element being a binary integer represents nonconvexity. This additional binary integer constraint explains why convex technologies contain their nonconvex counterparts: $T^{\Lambda^{\text{C}}, \Gamma} \supseteq T^{\Lambda^{\text{NC}}, \Gamma}$. The scaling parameter δ is free under CRS, smaller than or equal to 1 or larger than or equal to 1 under NIRS and NDRS respectively, and fixed at 1 under VRS.¹¹ This unified formulation has a pedagogical advantage in that it clearly separates the role of the various assumptions in the formulation of the technology.

Computing the input distance function (1) relative to convex technologies in (5) requires solving a nonlinear programming (NLP) problem for each evaluated observation. This NLP can be easily transposed into the familiar linear programming (LP) problem that are found

¹¹ Nonparametric deterministic frontier estimators that minimally extrapolate the data subject to some maintained axioms are labeled Data Envelopment Analysis (DEA) models in contrast to different (e.g., parametric) methodologies. The same moniker is also used to distinguish these convex models from the nonconvex ones in (5) that are sometimes labeled Free Disposal Hull (FDH) technologies. This use of monikers to denote both general methodologies and specific models is confusing.

in the literature (see Färe (1988) or Hackman (2008)).¹² For the nonconvex technologies, nonlinear mixed integer programs must be solved, but an implicit enumeration strategy is available.

Using the unified algebraic representation of the technology $T^{\Lambda, \Gamma}$ determined by (5), the cost function (2) can now be rewritten as

$$C(y, w \mid T^{\Lambda, \Gamma}) = \min \left\{ wx : x \geq \sum_{k=1}^K x_k \delta z_k, y \leq \sum_{k=1}^K y_k \delta z_k, z \in \Lambda, \delta \in \Gamma \right\}. \quad (6)$$

Since convex technologies contain their nonconvex counterparts, convex cost functions are smaller than or equal to their nonconvex counterparts. In the remainder, this cost function notation is simplified to highlight either the nature of convexity or the nature of returns to scale. This Eq. (6) leads directly to the following lemma:

Lemma 3.1 *For a given output level $y \in \mathbb{R}_+^m$ and a price vector $w \in \mathbb{R}_+^n$, $C(y, w \mid \text{CRS}) = \min\{C(y, w \mid \text{NIRS}), C(y, w \mid \text{NDRS})\}$.*

Proof This result follows directly from the fact that $\Gamma^{\text{CRS}} = \Gamma^{\text{NIRS}} \cup \Gamma^{\text{NDRS}}$. □

Computing these cost functions relative to convex nonparametric technologies involves solving an LP per observation (see Färe (1988) or Hackman (2008)). For the cost functions relative to the nonconvex technologies, again implicit enumeration algorithms have been obtained. Note again that the VRS nonconvex cost function is equivalent to WACM.

Recalling from Proposition 2.1 that costs evaluated on nonconvex technologies are higher or equal to costs evaluated on convex technologies (except under CRS and a single output), this relation implies that if the true technology is nonconvex, then a cost function computed on a convex technology is inconsistent and always yields a downwardly biased estimate. The reverse need not hold true: for a truly convex technology a cost function computed on a nonconvex technology remains consistent and still asymptotically converges to the one computed on a convex technology albeit at a slower rate of convergence (see Simar and Wilson (2000, p. 56) for these statements in the context of technologies).

While the above approach can be considered deterministic, since all deviations from the frontier are attributed to inefficiencies, several proposals are available in the literature to develop stochastic nonparametric frontier estimators (see Olesen and Petersen 2016 for a balanced survey). In view of the Duhem-Quine hypothesis that the empirical testing of any scientific hypothesis requires auxiliary assumptions and that in case of falsification it is difficult to single out one hypothesis responsible for its failure, we keep the amount of auxiliary assumptions to a strict minimum by sticking to the simplest of theoretical frameworks in which convexity can be tested. Therefore, our estimates combine inefficiency and measurement error.¹³

¹² Substituting $t_k = \delta z_k$ in (5), one rewrites the sum constraint on the activity vector. Realizing that the constraints on the scaling factor are integrated into the latter sum constraint, the LP appears.

¹³ Badunenko et al. (2012) show that the reliability of efficiency scores of nonparametric, deterministic convex frontier estimation remains excellent when the ratio of the variation in efficiency to the variation in noise is low. Assuming this result would also be confirmed for nonconvex similar estimators and would also hold for the unexplored case of the cost function, this reinforces our argument to opt for a simple deterministic nonparametric frontier framework to test for convexity.

3.2 Sections of cost functions in the outputs

The reconstruction and visualization of production frontiers has been the subject of a rather limited number of contributions (e.g., Hackman et al. (1994) or Hackman (2008, Ch. 10)). Since traditional production possibility sets are convex polyhedra, one can enumerate their facets to reconstruct the boundaries of the set. A two-dimensional projection is then defined relative to a particular point in the technology. Krivonozhko et al. (2004) present a family of parametric optimization methods to construct an intersection of the multidimensional frontier with a two-dimensional plane determined by any pair of given directions.

In a similar way, for a given observation a section of a cost function along one particular output dimension with a grid of 10,000 points within the empirical range of the sample for the output selected can be computed using simple parametric programming methods. We also numerically approximate for the first time in the empirical literature a count of the number of change points spanning these piecewise linear cost functions. Sections in cost-output space are -to our knowledge- new in the literature: the algorithm is described in the Appendix A in Electronic supplementary material.

3.3 Characterizing returns to scale and economies of scale

Several methods have been proposed in the literature to obtain qualitative information regarding global returns to scale. Since these methods are unsuitable for nonconvex technologies, Kerstens and Vanden Eeckaut (1999, Proposition 2) generalize an existing goodness-of-fit method to suit all technologies. With the inclusion by Podinovski (2004b) of a fourth returns to scale case only relevant for nonconvex technologies, the following proposition summarizes this method.

Proposition 3.1 (Cesaroni, Kerstens, and Van de Woestyne (2017b, Proposition 3.3)) *Technology $T^{\Delta, VRS}$ is globally characterized at the optimal projected point in input direction (δ^*x, y) with $\delta^* = E_i(x, y | VRS)$ as follows:*

- (a) $CRS \Leftrightarrow E_i(x, y | NIRS) = E_i(x, y | NDRS) = E_i(x, y | VRS)$;
- (b) $IRS \Leftrightarrow E_i(x, y | NIRS) < E_i(x, y | NDRS) \leq E_i(x, y | VRS)$;
- (c) $DRS \Leftrightarrow E_i(x, y | NDRS) < E_i(x, y | NIRS) \leq E_i(x, y | VRS)$;
- (d) $SCRS \Leftrightarrow E_i(x, y | NIRS) = E_i(x, y | NDRS) < E_i(x, y | VRS)$;

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. Note that (Podinovski 2004a) convincingly shows that global and local returns to scale need not coincide under nonconvexity: for instance, when moving from an IRS towards a CRS point average productivity increases monotonously on a convex technology, but this need not hold true for a nonconvex technology. Note that the characterization in Proposition 3.1 should be considered at the optimal projected point in the input direction. Therefore, if starting from an inefficient observation, then the corresponding characterization refers to that of this corresponding optimal projected point.

We now develop a goodness-of-fit method using various cost functions to determine different economies of scale cases in a nonconvex (non-differentiable) setting. To achieve this, we first introduce the related notions of minimal ray average cost and optimal scale

size in the following definition (following Chavas and Cox 1999 and especially Cesaroni and Giovannola 2015).

Definition 3.1 For a given output level $y \in \mathbb{R}_+^m$ and a price vector $w \in \mathbb{R}_+^n$, the minimal ray average cost (RAC) with respect to technology T is given by

$$\text{RAC}(y, w \mid T) = \min_{\lambda} \left\{ \frac{C(\lambda y, w \mid T)}{\lambda} : \lambda > 0 \right\}.$$

The optimal value of λ realizing RAC is referred to as the optimal scale size.

The following lemma provides some relations between RAC and the cost function:

Lemma 3.2 For a given output level $y \in \mathbb{R}_+^m$ and a price vector $w \in \mathbb{R}_+^n$:

- (a) $\text{RAC}(y, w \mid \text{CRS}) = C(y, w \mid \text{CRS})$;
- (b) $\text{RAC}(y, w \mid \text{VRS}) = C(y, w \mid \text{CRS})$;
- (c) $\text{RAC}(y, w \mid \text{CRS}) = \text{RAC}(y, w \mid \text{VRS}) = C(y, w \mid \text{CRS})$;
- (d) $\text{RAC}(y, w \mid T) \leq C(y, w \mid T)$.

Part (b) of Lemma 3.2 has been first developed in Chavas and Cox (1999). For convenience, the proof of Lemma 3.2 is in the Appendix B in Electronic Supplementary material. Using RAC, we can now define economies of scale of the cost function for the (true) VRS technology:

Definition 3.2 For a given output level $y \in \mathbb{R}_+^m$ and a price vector $w \in \mathbb{R}_+^n$, the following economies of scale are provided for $C(y, w \mid \text{VRS})$:

- (a) CES if $\text{RAC}(y, w \mid \text{VRS})$ is attained for at least the optimal scale size 1;
- (b) IES if $\text{RAC}(y, w \mid \text{VRS})$ is attained for one or more optimal scale sizes, all of which are strictly greater than 1;
- (c) DES if $\text{RAC}(y, w \mid \text{VRS})$ is attained for one or more optimal scale sizes, all of which are strictly smaller than 1;
- (d) SCES if $\text{RAC}(y, w \mid \text{VRS})$ is attained for multiple optimal scale sizes, some of which are strictly larger than 1 and others are strictly smaller than 1,

where CES, IES, DES and SCES stand for constant, increasing, decreasing and sub-constant economies of scale, respectively.

This approach is introduced in the literature by Cesaroni and Giovannola (2015, Section 5) and is equivalent to Cesaroni, Kerstens, and Van de Woestyne (2017a, Proposition 1). Note again that the fourth economies of scale case is only relevant for nonconvex cost functions. These types of economies of scale are characterized by the following practical proposition:

Proposition 3.2 For a given output level $y \in \mathbb{R}_+^m$ and a price vector $w \in \mathbb{R}_+^n$, the following characterizations for economies of scale of $C(y, w \mid \text{VRS})$ hold true:

- (a) CES if $C(y, w \mid \text{NIRS}) = C(y, w \mid \text{NDRS}) = C(y, w \mid \text{VRS})$;
- (b) IES if $C(y, w \mid \text{NIRS}) < C(y, w \mid \text{NDRS}) \leq C(y, w \mid \text{VRS})$;
- (c) DES if $C(y, w \mid \text{NDRS}) < C(y, w \mid \text{NIRS}) \leq C(y, w \mid \text{VRS})$;
- (d) SCES if $C(y, w \mid \text{NIRS}) = C(y, w \mid \text{NDRS}) < C(y, w \mid \text{VRS})$.

The proof of Proposition 3.2 is in the Appendix B in Electronic Supplementary material. Observe that the structure of Proposition 3.2 is very similar to Proposition 3.1 dealing with returns to scale in production correspondences. In fact, if only one input is assumed, then the following result holds true:

Lemma 3.3 *Consider a technology T with only one input and, hence, one input price $w \in \mathbb{R}_+$. For $(x, y) \in T$, it follows that*

$$\frac{C(y, w | T)}{w} = xE_i(x, y | T).$$

The proof of Lemma 3.3 is in the Appendix B in Electronic Supplementary material. Using this lemma, the correspondence between Propositions 3.1 and 3.2 follows directly in the case of a single input. Obviously, this relation no longer applies in the case of multiple inputs.

4 Description of the samples

We employ two secondary data sets for our empirical analysis from the Journal of Applied Econometrics Data Archive (<http://qed.econ.queensu.ca/jae/>). The first sample is based on 16 Chilean hydro-electric power generation plants observed on a monthly basis for several years (Atkinson and Dorfman 2009). Limiting ourselves to the observations for the single year 1997, we can safely ignore any technical change and specify an intertemporal frontier. This results in a total of 192 observations. There is one output quantity (electricity generated). There are also the prices and quantities of three inputs: labor, capital, and water. Except for the input capital, all remaining flow variables are expressed in physical units. Prices are in current Chilean pesos. Basic descriptive statistics for the inputs and the single output as well as more details on these data are available in Atkinson and Dorfman (2009).

The second sample is an unbalanced panel of three years (1984–1986) of French fruit producers based on annual accounting data collected in a survey (Ivaldi et al. 1996). Mainly two criteria were adapted to select the farms: (i) the production of apples must be larger than zero, and (ii) the productive acreage of the orchard must be at least five acres. As a technology, three aggregate inputs are combined to produce two outputs. The three inputs are: (i) capital (including land), (ii) labor, and (iii) materials. The two aggregate outputs are (i) the production of apples, and (ii) an aggregate of alternative products. Also input prices are available in French francs. Summary statistics for the 405 observations in total and details on the definitions of all variables are available in Appendix 2 in Ivaldi et al. (1996). Note that the short length of the panel (only three years) justifies the use of an intertemporal approach that ignores technical change.

This choice of data sets is driven by the desire to illustrate the effects of the convexity assumption on the cost function in both the single and the multiple output cases.

5 Empirical results

Recall that first we report the eventual differences between estimates based on convex and nonconvex cost functions at the sample level. Furthermore, these are illustrated with graphical sections relating costs to the outputs for some specific observations. Second, we also illustrate the eventual differences between the characterization of economies of scale and returns to scale for convex and nonconvex cost functions and technologies alike.

5.1 Cost frontier estimates: descriptive statistics and testing convexity

Table 2 reports some basic descriptive statistics on both the convex and nonconvex cost frontier estimates on both VRS and CRS technologies. Since CRS is a strong and unrealistic

Table 2 Nonconvex and convex cost function estimates: descriptive statistics

Costs	Nonconvex (NC)		Convex (C)		Δ w.r.t. NC (%)	
	VRS	CRS	VRS	CRS	VRS	CRS
Sample 1: Chilean hydro-power plants						
Average	10.67	6.22	8.45	6.22	20.76	0.00
Stand.Dev.	13.64	7.43	11.08	7.43	18.78	0.00
Minimum	2.49	0.04	2.49	0.04	0.00	0.00
Maximum	65.80	39.25	65.80	39.25	0.00	0.00
Li test [†]	VRS 12.18***		CRS 0.0			
Sample 2: French fruit producers						
Average	1160.91	683.06	718.84	511.51	38.08	25.11
Stand.Dev.	1730.08	880.89	1124.45	758.76	35.01	13.86
Minimum	150.11	13.15	150.11	8.51	0.00	35.29
Maximum	13448.40	6754.19	11815.70	6095.27	12.14	9.76
Li test [†]	VRS 12.29***		CRS 7.34***			

[†]Li test: critical values at 1% = 2.33 (***); 5% = 1.64 (**); 10% = 1.28 (*)

assumption, the computations of the cost frontier estimates relative to CRS technologies mainly serves the purpose of illustrating Proposition 2.1. The last two columns report the difference in terms of the nonconvex estimates (i.e., $(C(y, w | T^{NC}) - C(y, w | T^C))/C(y, w | T^{NC})$), just as in the second column of Table 1. Three conclusions jump out. First, nonconvex cost frontier estimates are on average substantially higher than their convex cost counterparts, while the VRS ones are again higher than the CRS ones. The last two columns indicate that the convex estimates are on average between 20.76% and 38.08% lower in the VRS case, and between 0.00% and 25.11% lower in the CRS case. The single output sample reveals a smaller gap than the multi-output case. Second, for the hydro-power plants both nonconvex and convex results are identical for the CRS case, because of the single output. Third, while the empirical range of the VRS case is identical in the single output case, by contrast the empirical range varies substantially for the multi-output fruit producers, apart from the minimum cost difference of 0.00% on the VRS technology. The latter is explained by the fact that for 8 fruit producers (1.98%) convex and nonconvex cost estimates coincide.

A detailed analysis of the last two columns can furthermore reveal us something on the impact of multiple outputs and CRS in explaining the divergence between nonconvex and convex costs. In the multi-output case, the total gap between nonconvex and convex costs amounts to 38.08%. The amount of 25.11% is due to the multi-output nature of production given CRS, while the residual 12.97% (= 38.08% - 25.11%) is due to nonconstant returns to scale. In the single output case, the total gap of 20.76% is due to nonconstant returns to scale, since the assumption of CRS would reduce the gap to 0% by definition. The differences between both data sets call for more research into the causes of the total cost gap.

To facilitate the appreciation of these differences in more detail we also plot kernel density estimates of all cost estimates related to the four frontier specifications per data set. Figure 1a and b represent the densities for most part of the observed range of cost estimates for the hydro-power plants and the fruit producers respectively.¹⁴ Figure 1a illustrates clearly:

¹⁴ To enhance comparability we use a common Sheather and Jones plug-in bandwidth for the convex and nonconvex data series to be compared.

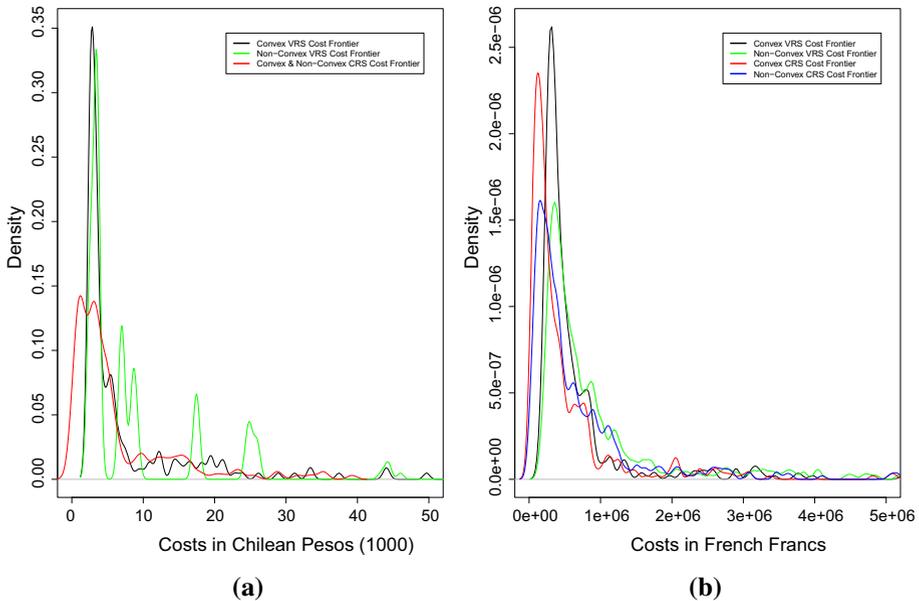


Fig. 1 Kernel density estimates of cost function estimates for **a** chilean hydro-power plants and **b** French fruit producers

(i) Convex and nonconvex VRS cost frontiers differ markedly in shape: the first is downwardly asymmetrical and almost uni-modal while the second is asymmetrical multi-modal. (ii) Convex and nonconvex CRS cost frontiers coincide: its distribution is about downwardly asymmetrical uni-modal. Figure 1b reveals downwardly asymmetrical densities for all frontiers and a markedly higher mode for the convex cases for a given returns to scale assumption.

Differences between the densities of these cost function estimates can be tested with a consistent nonparametric test of closeness between two unknown density functions that is asymptotically normal distributed under the corresponding null hypothesis. This Li test statistic -proposed by Li (1996) and refined by Fan and Ullah (1999)- has an important characteristic for our purpose: it is valid for both dependent and independent variables. Dependency is distinctive for frontier estimators (e.g., efficiency depends on sample size, among others). Furthermore, it imposes the quite mild conditions that both densities are bounded and continuous (see Fan and Ullah 1999, Section 2) and it works fine for moderate sample sizes above 50. The null hypothesis states the equality of both convex and nonconvex distributions of cost function estimates for a given returns to scale assumption. The bottom lines in upper and lower parts of Table 2 report the test statistics: apart from the identical series for CRS and a single output, the differences in densities between convex and nonconvex series are significant.

5.2 Sections of the cost function in the outputs

As a graphical illustration of the above sample level results, we plot convex and nonconvex cost functions for some specific observations starting from a grid of 10,000 points. Figure 2a plots cost estimated on VRS along the single output for hydro-power plant number 16 in July 1997. Figure 3a plots costs estimated on VRS along the first output dimension solely for

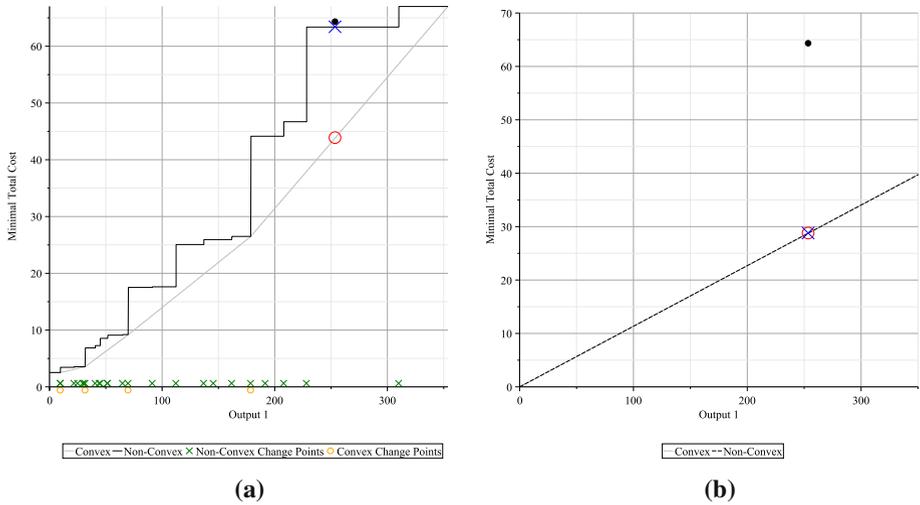


Fig. 2 Cost function in the single output for hydro-power plant 16 in July 1997 (a) VRS and (b) CRS

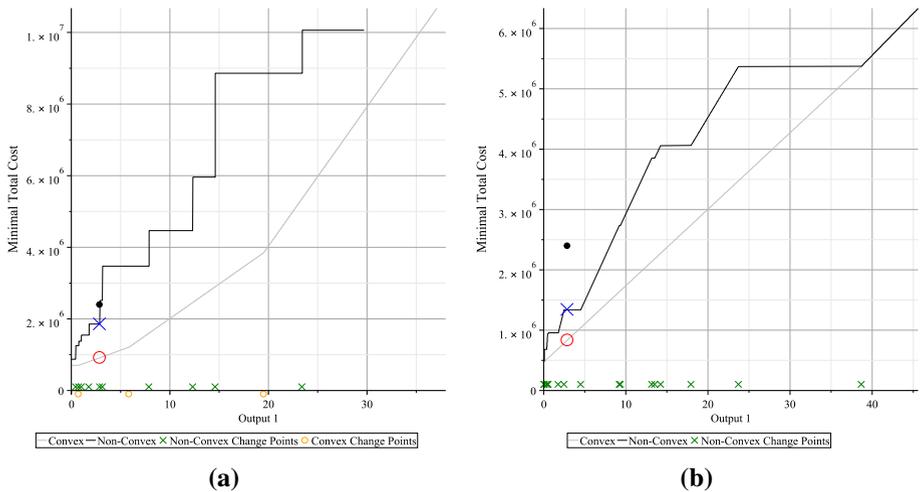


Fig. 3 Cost function in output 1 for fruit producer 19 in 1984 (a) VRS and (b) CRS

fruit producer number 19 in 1984. Figures 2b and 3b plot total costs for the same individual observations for the CRS case for the respective samples. Notice that for the single output hydro-power plants the convex and nonconvex CRS costs coincide.

Two key observations can be made. First, while in the single output case in Figure 2a, apart from the extremes both convex and nonconvex cost functions coincide several times within the empirical range of the data, both cost functions are entirely different within the empirical range of the data in the multiple output case plotted in Fig. 3a (again apart from the extremes).¹⁵ Second, for observations within the empirical range the nonconvex cost function

¹⁵ Indeed, in the single output case it may be intuitively clear that convex and nonconvex cost functions trace the same function at the extremes of the empirical range, but that in between the approximation depends on the

always offers a substantially closer fit compared to its convex counterpart. Representing the original observation by a bullet (●), the nonconvex and convex projection points on the cost frontiers are indicated by a times sign (×) and circle (○) respectively. One can observe that the difference between cost frontiers is huge in Fig. 2a for observation 16 in July 1997. This difference is slightly less pronounced for observation 19 in Fig. 3a.

For the CRS case, both projection points coincide in the single output case in Fig. 2b, but diverge in the multiple output case in Fig. 3b. Overall, this further casual evidence illustrates that convex and nonconvex costs may diverge substantially depending on the output level considered within the empirical range and the returns to scale assumption maintained.

These empirical observations may point to a difference in local and global fit between both cost functions: while the nonconvex cost function has both a good local and global fit, the convex cost function may sometimes have a good local fit but its global fit is poor. The local fit concerns the distance between the observation being analysed and its projection point on the frontier. This distance is always small for the nonconvex cost function, and it can be small or large for the convex cost function (just contrast Figs. 2a and 3a).

One possible indication of the global fit is a count of the number of change points spanning the piecewise linear cost functions and to assume that more change points indicate a better global fit.¹⁶ For the VRS cost functions depicted in Figs. 2a and 3a, we count on the one hand 4 and 24, and on the other hand 3 and 10 change points for the convex and nonconvex cases, respectively. For the CRS cost functions depicted in Figs. 2b and 3b, one finds on the one hand 0 and 0, and on the other hand 0 and 16 change points for the convex and nonconvex cases, respectively. The numerical procedure to determine these change points starts from the grid of 10,000 points. Clearly, the global fit of the nonconvex cost function is way better than that of its convex counterparts.¹⁷

To avoid the anecdotal nature of this evidence, Table 3 reports descriptive statistics on the number of change points for all observations in both samples. To reduce the computational burden, the numerical procedure determining the change points starts from a grid of 1000 points only. The number of change points are counted for each observation for the convex and nonconvex cases under both VRS and CRS using Algorithm A.2 provided in Appendix A in Electronic Supplementary material. Statistics are found in the first four columns (not counting the first header column). For the fruit producers, we report all change points along a section in the first output dimension (as depicted) as well as along a section in the second output dimension (which has not been depicted). Two observations are clear. First, under the VRS hypothesis the nonconvex sections show on average up to four times more change points compared to convexity (the CRS sections actually show even relatively more change points). Second, the situations depicted in Figs. 2a–3b clearly fall in line with these results and are not peculiar at all. This additional systematic evidence shows that convex and nonconvex cost frontier estimates are supported by a substantially different amount of observations determining these change points, depending on the returns to scale assumption.

precise structure of the sample. This intuition for the single output case is developed in the figures in Appendix C in Electronic Supplementary material by computing the models in Sect. 3.1 on a numerical example with two inputs and a single output. However, the situation is less clear in the multi-output case: convex and nonconvex cost functions need not have very much in common except at some of the extremes of the empirical range (see Appendix D in Electronic Supplementary material for figures computed on a numerical example with one input and two outputs).

¹⁶ As pointed out by a referee, there is a multitude of approaches to compute the difference between two functions and thus to assess global fit.

¹⁷ As pointed out by the same referee, to the best of our knowledge it remains an open question to which extent this better fit of the nonconvex cost function may lead to overfitting when predicting.

Table 3 Nonconvex and convex cost function sections: descriptive statistics on change points

Costs	Nonconvex (NC)		Convex (C)		Equal cost C & NC	
	VRS	CRS	VRS	CRS	VRS	CRS
Sample 1: Chilean hydro-power plants						
Average	19.44	0.00	4.13	0.00	6.16	0.00
Stand.Dev.	0.71	0.00	0.34	0.00	0.43	0.00
Minimum	19	0	4	0	6	0
Maximum	21	0	5	0	8	0
Sample 2: French fruit producers						
<i>Output 1</i>	<i>VRS</i>	<i>CRS</i>	<i>VRS</i>	<i>CRS</i>	<i>VRS</i>	<i>CRS</i>
Average	13.16	10.74	4.88	1.92	0.42	0.52
Stand.Dev.	2.94	2.94	0.95	0.28	0.57	0.55
Minimum	0	1	0	1	0	0
Maximum	18	19	8	2	2	2
<i>Output 2</i>	<i>VRS</i>	<i>CRS</i>	<i>VRS</i>	<i>CRS</i>	<i>VRS</i>	<i>CRS</i>
Average	5.95	9.75	3.69	0.98	1.59	0.81
Stand.Dev.	2.11	2.92	1.20	0.13	0.69	0.39
Minimum	0	1	0	0	0	0
Maximum	10	20	6	1	3	1

It is instructive to assess for how many of the change points in Table 3 convex and nonconvex cost estimates coincide. This is done as follows. For each observation, the union of the sets of all change points occurring in the convex and nonconvex case is considered. Obviously, this set only contains unique change points occurring either in the convex or nonconvex case. From these change points, those with the same optimal convex and nonconvex costs are counted. Statistics are reported in the last two columns for VRS and CRS, respectively.

Two conclusions emerge. First, the most favourable case is the hydro-power plants with a single output where on average 6.16 change points have identical costs. For the multiple output fruit producers these amounts are close to negligible. Second, the CRS case seems to be slightly worse than the VRS case, on average. Thus, it is clear that these results confirm the earlier results at the level of the individual observations: convex and nonconvex cost functions coincide only very rarely.

Our evidence on the hydro-power plants confirms the empirical evidence on nonconvexities in electricity generation costs in the engineering literature (e.g., Bjørndal and Jörnsten (2008)). The results for the fruit production farms reveal that, even in cases where a priori nonconvexities are not anticipated to be prominent, these still may play a role and influence cost function estimates. Obviously, part of the observed differences between convex and nonconvex cost function estimates may be due to stochastic errors or the fact that technical change did occur during the respective observation periods.¹⁸ However, it remains unclear how either stochastic errors or technical change would have a differential impact on convex versus nonconvex cost functions and thereby affect the observed differences.

¹⁸ Results in, e.g., Ray and Kim (1995) learn that accounting for nonregressive technical change increases technical efficiency, thereby improving the empirical fit. However, their panel data set covers a long time period, while ours have a short time dimension.

Table 4 Returns to scale and economies of scale: convex (C) vs. nonconvex (NC) models

Production	Chilean hydro-power plants (%)				French fruit producers (%)			
	C-IRS	C-CRS	C-DRS	Total	C-IRS	C-CRS	C-DRS	Total
NC-IRS	63.54	0.00	6.77	70.31	74.07	0.00	0.00	74.07
NC-CRS	11.46	2.60	2.60	16.67	9.38	1.48	1.98	12.84
NC-DRS	1.04	0.00	11.98	13.02	6.91	0.25	5.93	13.09
Total	76.04	2.60	21.35	100.00	90.37	1.73	7.90	100.00
Agreement				78.13				81.48
Disagreement				21.88				18.52
Cost	C-IES	C-CES	C-DES	Total	C-IES	C-CES	C-DES	Total
NC-IES	46.35	5.21	0.00	51.56	73.83	0.00	0.00	73.83
NC-CES	0.00	0.52	0.00	0.52	1.73	0.25	0.00	1.98
NC-DES	21.88	3.65	22.40	47.92	17.78	0.00	6.42	24.20
Total	68.23	9.38	22.40	100.00	93.33	0.25	6.42	100.00
Agreement				69.27				80.49
Disagreement				30.73				19.51

5.3 Returns to scale and economies of scale results

Next, it is important to verify whether there exist any differences in the determination of returns to scale and economies of scale for individual observations when applying Propositions 3.1 and 3.2. Per data set and per production and cost method, we report in Table 4 the percentages of observations for which the returns and economies to scale classification coincides, as well as the ones for which these classifications diverge. For each of the four parts in Table 4, starting from the nonconvex case on the left we see the impact of moving to a convex world in the columns to the right. In each part, row totals in the last column and column totals in the last row yield the returns and economies to scale for the nonconvex and the convex cases respectively. Furthermore, we report the percentage of diagonal (labeled “Agreement”) and off-diagonal (labeled “Disagreement”) elements in the total reflecting the consensus and the divergences among the results.

Focusing on the row totals in the last column and column totals in the last row in Table 4 per sample for both a production- and cost-based analysis yields the following conclusions. First, the majority of hydro-power plant and fruit producer observations are subjected to IRS. One qualification is that the nonconvex cost approach for the hydro-power plants in fact indicates about an equal amount of IRS and DRS. For both the hydro-power plants and the fruit producers the nonconvex cost approach reveals a larger share of observations subject to DRS compared to the production-based analysis. Second, there are more observations with CRS under nonconvexity, except for the cost approach applied to the hydro-power plants. Despite the costs being equal under CRS and a single output, perhaps surprisingly the determination of global returns to scale leads to more observations being characterised as operating under CES under a convex relative to a nonconvex technology.

We make two further observations on the results in Table 4. First, focusing on the diagonal elements of this table, consensus on the classification varies between 69.27% and 78.13% for the hydro-power plants. This leaves a wide to modest margin of conflict. The consensus is more stable for the fruit producers across production- and cost-based analysis. Moreover, for

both samples consensus is highest for the production approach. Second, turning to conflicting information, the extreme case of conflict among the off-diagonal elements is obviously the switch from increasing returns (economies) to scale to decreasing returns (diseconomies) to scale, or the reverse. For the hydro-power plants, this varies from an almost negligible 7.81% ($= 1.04\% + 6.77\%$) to an impressive 21.88% ($= 21.88\% + 0.00\%$) of cases for the production and the cost approaches respectively. The amount of conflict among the fruit producers is also quite substantial: it ranges between a mild 6.91% in production to a more pronounced 17.78% in the cost approach.

6 Conclusions

The main focus of this paper has been empirical rather than methodological: is convexification of nonparametric cost function estimates harmless, or does it create a bias. We think the global answer to this question has been that convexification is harmful.

Using data on Chilean hydro-power plants as well as French fruit producers, this contribution is the first to empirically illustrate the differences between convex and nonconvex cost frontier estimates and their resulting distributions, and to explore the effects on returns to scale and economies of scale. For these specific samples, convex cost functions may yield potentially downwardly biased estimates in the order between 0.00 % and 38.08 % depending on returns to scale hypothesis and the single vs. multi-output case. While these bias estimates may seem high, these numbers are in line with the scant evidence available in the literature: Table 1 even reports substantially higher biases. Not only are the resulting distributions of the cost estimates different at a statistically significant level, some sections of costs along a single output dimension have served to further illustrate that these differences can be huge for individual observations situated within the empirical range (not at its extremes). The sample of hydro-power plants also illustrates the result that convex and nonconvex cost functions coincide under CRS and a single output. A more systematic exploration of the number of change points supporting these empirical sections of costs along an output dimension reveals that nonconvex sections are spanned by substantially more points than convex sections.

Furthermore, differences between the characterization of economies of scale and returns to scale based on convex and nonconvex technology and cost function estimations have been highlighted. This characterization of both economies of scale and returns to scale for individual observations turns out to be conditioned by convexity in a nonnegligible way. In particular, for both samples consensus on returns to scale and economies of scale is highest for the production approach.

If cost estimates and the characterization of economies of scale and returns to scale for individual observations turn out to be conditioned by the convexity assumption, then this may have important consequences. For one, investment decisions to increase or decrease the scale of operations based on economies of scale or returns to scale information could be responding to the wrong signals. For another, some capacity notions are very closely linked with the notion of scale economies and therefore would suffer from the same problem. For instance, economic capacity as defined by Klein (1960), among others, considers the outputs determined by the minimum of the long run average total costs as a reference to determine practical capacity utilization ratios. The study by Kerstens et al. (2019) provides some first evidence that convex and nonconvex capacity concepts tend to differ. Equally so, it is conceivable that the measurement of economies of scope and economies of specialisation is affected by the impact of convexity on the cost function (see, e.g., the framework developed

by Oude Lansink et al. (2015)). Thus, our findings have potentially important consequences for investment decisions, definitions of capacity utilization notions, and other key economic notions.

Furthermore, the use of convex cost estimates in price cap regulation may well be too harsh in sectors where nonconvexities play a major role (e.g., electricity generation and distribution). Given the better fit of nonconvex cost estimates, for the resulting efficiency requirements to have some bite it may well become necessary to either increase the geographical scope of samples, or increase the frequency of observations. The former solution is sometimes developed by international coordination in regulatory practices (see, e.g., Estache et al. (2004)). Conditional upon the availability of engineering models, one can also substitute or complement existing data with so-called pseudo-data to obtain larger sample sizes (see Griffin (1979)).¹⁹ Such engineering benchmarks or normative models (based on engineering production and cost models) are in fact used in several countries (e.g., Chile, Spain and Sweden).

Therefore, it seems important to empirically explore these differences between estimates based on convex and nonconvex technologies and cost functions further in even greater detail (e.g., also focusing on economies of scope, the impact on mergers and acquisitions, the effect on marginal cost relationships, etc.). In addition, it may be envisioned to set up simulation studies to further explore the impact of the convexity assumption on technologies and cost functions alike under the assumption that the true technology and cost function are either convex or nonconvex. In conclusion, even though theoretically the impact of convexity has been known since some time, it seems to be important to further explore the effects of convexity on key economic value relations in practice. Some first evidence has been provided that the impact is nonnegligible and that convexification is not harmless for the cost function.

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 (e.g., the potential differences between global and local returns to scale (see Podinovski (2004a)), new standards may need to be established. For instance, Scarf (1986) associates to each feasible point a neighbourhood system to test for global optimality. This lack of standards to report nonconvex results as well as the need to go beyond traditional convex optimisation that is often considered a cornerstone for economic analysis may well contribute to its negligence.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10479-021-04069-1>.

References

- Afriat, S. (1972). Efficiency estimation of production functions. *International Economic Review*, 13(3), 568–598.
- Alam, I., & Sickles, R. (2000). Time series analysis of deregulatory dynamics and technical efficiency: the case of the US airline industry. *International Economic Review*, 41(1), 203–218.
- Allahverdi, A., Ng, C., Cheng, T., & Kovalyov, M. (2008). A survey of scheduling problems with setup times or costs. *European Journal of Operational Research*, 187(3), 985–1032.
- Ang, F., Mortimer, S., Areal, F., & Tiffin, R. (2018). On the opportunity cost of crop diversification. *Journal of Agricultural Economics*, 69(3), 794–814.

¹⁹ This proposal related to the engineering economics literature is not without criticism (see Maddala and Roberts (1981)). More recently, Preckel and Hertel (1988) propose a linear program summary functions method (LPSF) that is equivalent to the pseudodata approach with an infinite sample.

- Atkinson, S., & Dorfman, J. (2009). Feasible estimation of firm-specific allocative inefficiency through Bayesian numerical methods. *Journal of Applied Econometrics*, 24(4), 675–697.
- Badunenko, O., Henderson, D., & Kumbhakar, S. (2012). When, where and how to perform efficiency estimation. *Journal of the Royal Statistical Society*, 175A(4), 863–892.
- Balaguer-Coll, M., Prior, D., & Tortosa-Ausina, E. (2007). On the determinants of local government performance: a two-stage nonparametric approach. *European Economic Review*, 51(2), 425–451.
- Barnett, W. (2002). Tastes and technology: curvature is not sufficient for regularity. *Journal of Econometrics*, 108(1), 199–202.
- Bjørndal, M., & Jörnsten, K. (2008). Equilibrium prices supported by dual price functions in markets with non-convexities. *European Journal of Operational Research*, 190(3), 768–789.
- Bogetoft, P., Tama, J., & Tind, J. (2000). Convex input and output projections of nonconvex production possibility sets. *Management Science*, 46(6), 858–869.
- Briec, W., Kerstens, K., & Vanden Eeckaut, P. (2004). Non-convex technologies and cost functions: definitions, duality and nonparametric tests of convexity. *Journal of Economics*, 81(2), 155–192.
- Brokken, R. (1977). The case of a queer isoquant: increasing marginal rates of substitution of grain for roughage in cattle finishing. *Western Journal of Agricultural Economics*, 1(1), 221–224.
- Budish, E., Cramton, P., & Shim, J. (2014). Implementation details for frequent batch auctions: slowing down markets to the blink of an eye. *American Economic Review*, 104(5), 418–424.
- Cesaroni, G., & Giovannola, D. (2015). Average-cost efficiency and optimal scale sizes in non-parametric analysis. *European Journal of Operational Research*, 242(1), 121–133.
- Cesaroni, G., Kerstens, K., & Van de Woestyne, I. (2017a). Estimating scale economies in non-convex production models. *Journal of the Operational Research Society*, 68(11), 1442–1451.
- Cesaroni, G., Kerstens, K., & Van de Woestyne, I. (2017b). Global and local scale characteristics in convex and nonconvex nonparametric technologies: a first empirical exploration. *European Journal of Operational Research*, 259(2), 576–586.
- Chavas, J.-P., & Cox, T. (1999). A generalized distance function and the analysis of production efficiency. *Southern Economic Journal*, 66(2), 294–318.
- Copeland, A., & Hall, G. (2011). The response of prices, sales, and output to temporary changes in demand. *Journal of Applied Econometrics*, 26(2), 232–269.
- Coviello, D., Ichino, A., & Persico, N. (2014). Time allocation and task juggling. *American Economic Review*, 104(2), 609–623.
- Cummins, D., & Zi, H. (1998). Comparison of frontier efficiency methods: an application to the U.S. life insurance industry. *Journal of Productivity Analysis*, 10(2), 131–152.
- De Borger, B., & Kerstens, K. (1996). Cost efficiency of Belgian local governments: a comparative analysis of FDH, DEA, and econometric approaches. *Regional Science and Urban Economics*, 26(2), 145–170.
- Diewert, W., & Parkan, C. (1983). Linear programming test of regularity conditions for production functions. In W. Eichhorn, K. Neumann, & R. Shephard (Eds.), *Quantitative Studies on Production and Prices* (pp. 131–158). Würzburg: Physica-Verlag.
- Eaton, B., & Lipsey, R. (1997). *On the foundations of monopolistic competition and economic geography: the selected essays of B. Curtis Eaton and Richard G. Lipsey*. Cheltenham: Edward Elgar.
- Estache, A., Rossi, M., & Ruzzier, C. (2004). The case for international coordination of electricity regulation: evidence from the measurement of efficiency in South America. *Journal of Regulatory Economics*, 25(3), 271–295.
- Fan, Y., & Ullah, A. (1999). On goodness-of-fit tests for weakly dependent processes using kernel method. *Journal of Nonparametric Statistics*, 11(1), 337–360.
- Färe, R. (1988). *Fundamentals of production theory*. Berlin: Springer.
- Farrell, M. (1959). The convexity assumption in the theory of competitive markets. *Journal of Political Economy*, 67(4), 377–391.
- Fuss, M., McFadden, D., & Mundlak, Y. (1978). A survey of functional forms in the economic analysis of production. In M. Fuss & D. McFadden (Eds.), *Production economics: a dual approach to theory and applications* (Vol. 1, pp. 219–268). Amsterdam: North-Holland.
- Garbaccio, R., Hermalin, B., & Wallace, N. (1994). A comparison of nonparametric methods to measure efficiency in the savings and loan industry. *Real Estate Economics*, 22(1), 169–193.
- Grifell-Tatjé, E., & Kerstens, K. (2008). Incentive regulation and the role of convexity in benchmarking electricity distribution: economists versus engineers. *Annals of Public and Cooperative Economics*, 79(2), 227–248.
- Griffin, J. (1979). Statistical cost analysis revisited. *Quarterly Journal of Economics*, 93(1), 107–129.
- Hackman, S. (2008). *Production economics: integrating the microeconomic and engineering perspectives*. Berlin: Springer.

- Hackman, S., Passy, U., & Platzman, L. (1994). Explicit representation of the two-dimensional section of a production possibility set. *Journal of Productivity Analysis*, 5(2), 161–170.
- Hasenkamp, G. (1976). A study of multiple-output production functions. *Journal of Econometrics*, 4(3), 253–262.
- Hung, N., Le Van, C., & Michel, P. (2009). Non-convex aggregate technology and optimal economic growth. *Economic Theory*, 40(3), 457–471.
- Ivaldi, M., Ladoux, N., Ossard, H., & Simioni, M. (1996). Comparing Fourier and translog specifications of multiproduct technology: evidence from an incomplete panel of french farmers. *Journal of Applied Econometrics*, 11(6), 649–668.
- Izadi, H., Johnes, G., Oskrochi, R., & Crouchley, R. (2002). Stochastic frontier estimation of a CES cost function: the case of higher education in Britain. *Economics of Education Review*, 21(1), 63–71.
- Jacobsen, S. (1970). Production correspondences. *Econometrica*, 38(5), 754–771.
- Kerstens, K., & Managi, S. (2012). Total factor productivity growth and convergence in the petroleum industry: empirical analysis testing for convexity. *International Journal of Production Economics*, 139(1), 196–206.
- Kerstens, K., Sadeghi, J., & Van de Woestyne, I. (2019). Convex and nonconvex input-oriented technical and economic capacity measures: an empirical comparison. *European Journal of Operational Research*, 276(2), 699–709.
- Kerstens, K., & Vanden Eeckaut, P. (1999). Estimating returns to scale using nonparametric deterministic technologies: a new method based on goodness-of-fit. *European Journal of Operational Research*, 113(1), 206–214.
- Klein, L. (1960). Some theoretical issues in the measurement of capacity. *Econometrica*, 28(2), 272–286.
- Krivonozhko, V., Utikin, O., Volodin, A., Sablin, I., & Patrin, M. (2004). Constructions of economic functions and calculations of marginal rates in DEA using parametric optimization methods. *Journal of the Operational Research Society*, 55(10), 1049–1058.
- Kumbhakar, S., & Lovell, C. (2000). *Stochastic frontier analysis: an econometric approach*. Cambridge: Cambridge University Press.
- Li, Q. (1996). Nonparametric testing of closeness between two unknown distribution functions. *Econometric Reviews*, 15(1), 261–274.
- Li, S.-K. (2019). A nonparametric test of quasiconcave production function with variable returns to scale. *Journal of Mathematical Economics*, 82, 160–170.
- Maddala, G., & Roberts, R. (1981). Statistical cost analysis re-revisited: comment. *Quarterly Journal of Economics*, 96(1), 177–182.
- Mas-Colell, A., Whinston, A., & Green, J. (1995). *Microeconomic theory*. Oxford: Oxford University Press.
- Narbón-Perpiñá, I., Balaguer-Coll, M., Petrović, M., & Tortosa-Ausina, E. (2020). Which estimator to measure local governments' cost efficiency? the case of Spanish municipalities. *SERIEs*, 11(1), 51–82.
- Narbón-Perpiñá, I., Balaguer-Coll, M., & Tortosa-Ausina, E. (2019). Evaluating local government performance in times of crisis. *Local Government Studies*, 45(1), 64–100.
- Olesen, O., & Petersen, N. (2016). Stochastic data envelopment analysis-a review. *European Journal of Operational Research*, 251(1), 2–21.
- Oude Lansink, A., Stefanou, S., & Kapelko, M. (2015). The impact of inefficiency on diversification. *Journal of Productivity Analysis*, 44(2), 189–198.
- Podinovski, V. (2004a). Efficiency and global scale characteristics on the no free lunch assumption only. *Journal of Productivity Analysis*, 22(3), 227–257.
- Podinovski, V. (2004b). Local and global returns to scale in performance measurement. *Journal of the Operational Research Society*, 55(2), 170–178.
- Preckel, P., & Hertel, T. (1988). Approximating linear programs with summary functions: pseudodata with an infinite sample. *American Journal of Agricultural Economics*, 70(2), 397–402.
- Ray, S. (1997). Weak axiom of cost dominance: a nonparametric test of cost efficiency without input quantity data. *Journal of Productivity Analysis*, 8(2), 151–165.
- Ray, S. (2004). *Data envelopment analysis: theory and techniques for economics and operations research*. Cambridge: Cambridge University Press.
- Ray, S., & Kim, H. (1995). Cost efficiency in the US steel industry: a nonparametric analysis using data envelopment analysis. *European Journal of Operational Research*, 80(3), 654–671.
- Ray, S., & Mukherjee, K. (1995). Comparing parametric and nonparametric measures of efficiency: a reexamination of the Christensen-Greene data. *Journal of Quantitative Economics*, 11(1), 155–168.
- Resti, A. (1997). Evaluating the cost-efficiency of the Italian banking system: What can be learned from the joint application of parametric and non-parametric techniques. *Journal of Banking & Finance*, 21(2), 221–250.

- Romer, P. (1990). Are nonconvexities important for understanding growth? *American Economic Review*, 80(2), 97–103.
- Sauer, J. (2006). Economic theory and econometric practice: parametric efficiency analysis. *Empirical Economics*, 31(4), 1061–1087.
- Savani, R. (2012). High-frequency trading: the faster, the better? *IEEE Intelligent Systems*, 27(4), 70–74.
- Scarf, H. (1986). Testing for optimality in the absence of convexity. In W. Heller, R. Starr, & S. Starrett (Eds.), *Social choice and public decision making: Essays in honor of Kenneth J. Arrow* (Vol. I, pp. 117–134). Cambridge: Cambridge University Press.
- Scarf, H. (1994). The allocation of resources in the presence of indivisibilities. *Journal of Economic Perspectives*, 8(4), 111–128.
- Shephard, R. (1970). *Theory of cost and production functions*. Princeton: Princeton University Press.
- Simar, L., & Wilson, P. (2000). Statistical inference in nonparametric frontier models: the state of the art. *Journal of Productivity Analysis*, 13(1), 49–78.
- Varian, H. (1984). The nonparametric approach to production analysis. *Econometrica*, 52(3), 579–597.
- Viton, P. (2007). Cost efficiency in US air carrier operations, 1970–1984: a comparative study. *International Journal of Transport Economics*, 34(3), 369–401.
- Wibe, S. (1984). Engineering production functions: a survey. *Economica*, 51(204), 401–411.
- Winter, S. (2008). Scaling heuristics shape technology! should economic theory take notice? *Industrial and Corporate Change*, 17(3), 513–531.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.