May 2017

WORKING PAPER SERIES
2017-EQM-03

How Harmless is Convexification for Nonparametric Cost Function Estimates?

Kristiaan Kerstens
CNRS (LEM-UMR 9221) and IESEG School of Management

Ignace Van de Woestyne
KU Leuven, Research Unit MEES, BrusselsBelgium

IESEG School of Management
Lille Catholic University
3, rue de la Digue
F-59000 Lille
www.ieseg.fr

How Harmless is Convexification for Nonparametric Cost Function Estimates?*

Kristiaan Kerstens[†], Ignace Van de Woestyne[‡]

May 2017

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.

DRAFT: DO NOT QUOTE OR CITE WITHOUT PERMISSION.
Warning: Please print the figures in color for evaluation purposes.

JEL CODES: D24

KEYWORDS: Economics; Technology; Cost analysis

*We acknowledge the most helpful comments of G. Cesaroni, P. Kerstens, F. Maier-Rigaud and M. Vardanyan on earlier versions. We thank seminar participants at DIW (Berlin) and Aalto University (Helsinki) for useful comments on an early draft. The usual disclaimer applies.

†Corresponding author: CNRS-LEM (UMR 9221), IESEG School of Management, 3 rue de la Digue, F-59000 Lille, France, Tel: +33 320545892 (switchboard), Fax: +33 320574855, `k.kerstens@ieseg.fr`

‡KU Leuven, Research Unit MEES, Brussels, Belgium, `ignace.vandewoestyne@kuleuven.be`

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, Le Van, and Michel (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 states: "the essence of economies of scale in production is the presence of large and signicant 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 specic 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. It is our basic contention that the eventual impact of the convexity axiom on production and especially cost estimates has rarely been explicitly tested in the economic literature. Therefore, convexity can only be maintained if there is sufficient 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 as long as information on how well convex cost functions approximate nonconvex ones is largely missing.

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, 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. Third, these differences are also illustrated with graphical sections relating convex and nonconvex cost estimates to the outputs for individual observations. To the best of our knowledge, the combination of these aims is unique and did 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, McFadden, and Mundlak (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 (ii) 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, in Section 4 we introduce the two data sets employed in the empirical application. Following the structure of Section 3, Section 5 presents all empirical results in detail. Section 6 concludes.

2 Technology and Cost Functions: Role of Convexity

2.1 Technology and Cost Function: Definitions and Duality

We start by defining technology and some basic notation. 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: $L(y) = \{x \in \mathbb{R}^n_+ : (x, y) \in T\}$.

Throughout this contribution, 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.

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., Hackman (2008), Ray (2004) or Varian (1984) 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 (e.g., Sauer (2006) for a review of several studies). A sharp criticism of the constant returns to scale 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 Subsection 3.3). Finally, several reasons to doubt on the relevance of convexity are discussed in Subsections 2.2 and 2.3.

The input distance function offers a complete characterization of 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., 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.

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. 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 technology or a convex input set (see (Hackman, 2008, Ch. 7) for details). Briec, Kerstens, and Vanden Eeckaut (2004) are among the first to establish a local duality result between nonconvex technologies obeying different scaling laws and the corresponding nonconvex cost functions.[1]

[1] Earlier, already First, Hackman, and Passy (1993) establish duality for nonconvex technologies in a hierarchical organization sharing a common budget constraint.

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 nondecreasing and convex (nonconvex) in the outputs if and only if convexity of technology is assumed (rejected) (e.g., Jacobsen (1970) or Shephard (1970; 1974)). In a similar vein, Briec, Kerstens, and Vanden Eeckaut (2004) establish that cost functions estimated on convex technologies yield lower or equal cost estimates compared to cost functions estimated on nonconvex technologies[2]. Both types of cost functions are only identical for the single output and constant returns to scale case. These properties have been ignored in the large majority of 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.[3] 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, Ng, Cheng, and Kovalyov (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, Ichino, and Persico (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., see Savani (2012)).[4] Thus, perfect time divisibility seems a questionable assumption that ideally needs empirical validation. While all reasons commonly advanced to explain nonconvexities in production are features of the real economy (including indivisibilities in space), 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

[2]Ray (2004) shows that the latter cost function is just equivalent to the Weak Axiom of Cost Minimisation (WACM) defined by Varian (1984)

[3]One can also object that the time divisibility interpretation makes time enter into what essentially is a static theory of production. In this view, arguments related to time can only enter into a dynamic theory of production (see also Hackman (2008) for the latter).

[4]Budish, Cramton, and Shim (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.

nonconvexities exacerbate this time indivisibility and it remains 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 legitimation 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 no sense. Obviously, a nonconvex or a convex input set yield the same cost function (see Varian (1984)),[5] 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 constant returns to scale. This constant returns to scale 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).[6]

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

[5] Just as nonconvex or a convex technology yield 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).

[6] If one is willing to accept constant returns to scale, 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).

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.

First, the empirical evidence on engineering cost estimates available on some sectors (e.g., electricity generation) seems to point 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 have been amply documented in, e.g., Bjørndal and Jörnsten (2008) and Park, Jeong, Shin, and Lee (2010).[7] However, comparisons with convex estimates are unknown to us. Seemingly, nonconvex models are deemed a necessity and such comparisons have no use for engineers.

Second, in the economic literature one should 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) provide a recent example analyzing 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[8] 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 that 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 exis-

[7] One could cite tens of similar articles from specialised journals (e.g., *IEEE Transactions on Power Systems*).

[8] Recently, Diewert and Fox (2014) propose an alternative primal productivity index based on the same nonconvex technologies.

tence 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)).[9]

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, Johnes, Oskrochi, and Crouchley (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.[10]

A to the best of our knowledge complete list of cost studies in the economic literature somehow testing the convexity assumption is provided in Table 1. While the first column contains the author names, the second column lists the percentage difference between the convex ($C^C(y, w)$) and nonconvex ($C^{NC}(y, w)$) costs relative to nonconvex costs, while the third column adds some qualifying remarks.[11] All studies have multiple inputs producing multiple outputs, unless otherwise indicated in the third column. Note that only few studies mention 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 Section 3.1). First, Balaguer-Coll, Prior, and Tortosa-Ausina (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 10000) to 13.38% (population above 20000). Briec, Kerstens, and Vanden Eeckaut (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

[9]Müser and Dyckhoff (2017) recently employ an engineering model describing the process of quality splitting in waste incineration plants which reveals another source of nonconvexity in production.

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

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

coincide). The Cummins and Zi (1998) study focuses on 445 US life insurance companies over a balanced panel covering the period 1988–1992 and reveals 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, Hermalin, and Wallace (1994) report results for the US Savings and Loan industry in 1986: they find a difference of 23.46% and 22.37% for firms that remain solvent and insolvent, respectively, in the years 1988 to 1990.[12] 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. 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.[13] 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.

We end with two peculiar cases. First, car manufacturing costs are known to be nonconvex 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. Second, 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.

[12] 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%.

[13] The nonconvex results are presented graphically in the article. These numbers are based on supplementary information provided by the author.

Given that the vast majority of empirical approaches in production and cost analysis implicitly or explicitly maintains convexity of technology, one may conjecture that perhaps some 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 may well be nonconvex and cost function estimates on convex technologies may well be 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.

Table 1 about here

3 Empirical Methodology

3.1 Nonparametric Convex and Nonconvex 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, Kerstens, and Vanden Eeckaut (2004)). Some additional nonparametric, nonconvex production models are discussed in, e.g., Hackman (2008). The remainder of this subsection draws on Briec, Kerstens, and Vanden Eeckaut (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. 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.[14]

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 around in the

[14]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.

literature (see Hackman (2008)).[15] For the nonconvex technologies, nonlinear mixed integer programs must be solved, but an implicit enumeration strategy is available.[16]

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

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

In the remainder, this cost function notation is simplified to highlight either the nature of convexity or the nature of returns to scale. This equation (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 Hackman (2008)). For the cost functions relative to the nonconvex technologies, again implicit enumeration algorithms have been obtained.

It has been proven that costs evaluated on nonconvex technologies ($C(y, w \mid \text{NC})$) are higher or equal to costs evaluated on convex technologies ($C(y, w \mid \text{C})$)):

$$C(y, w \mid \text{NC})) \geq C(y, w \mid \text{C})). \tag{7}$$

Only in case of CRS and a single output: $C(y, w \mid \text{NC})) = C(y, w \mid \text{C}))$. This relation implies that if the true technology is nonconvex, then a cost function computed on a convex technology yields a downwardly biased estimate. The reverse need not hold: for a truly convex technology a cost function computed on a nonconvex technology asymptotically converges to the one computed on a convex technology.

This relation (7) can also be related to a property of the cost function in the outputs that seems often ignored: cost functions are nondecreasing in outputs and convex (nonconvex) in the outputs when the technology is convex (nonconvex) (see Jacobsen (1970, Proposition 5.2), or Shephard (1970, p. 227), Shephard (1974, p. 15)). The same reasoning applies to

[15] 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.

[16] Diewert and Fox (2014) derive a similar enumeration strategy for a primal productivity index.

the revenue function and all variations of the profit function, except of course the long run profit function where convexity does not matter for duality. Most advanced micro-economic textbooks seem to ignore this issue when, for instance, discussing the properties of the cost function (see Mas-Colell, Whinston, and Green (1995, p. 141)).

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. 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.[17]

3.2 Sections of Cost Functions

The reconstruction and visualization of production frontiers has been the subject of a rather limited number of contributions (e.g., Hackman, Passy, and Platzman (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, Utkin, Volodin, Sablin, and Patrin (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 10000 points within the empirical range of the sample for the output selected can be computed using simple parametric programming methods.

[17] Badunenko, Henderson, and Kumbhakar (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.

3.3 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. *Using $E_i(x, y \mid .)$ and conditional on the optimal efficient point, technology $T^{\Lambda,\text{VRS}}$ is globally characterized by:*

(a) CRS $\Leftrightarrow E_i(x, y \mid NIRS) = E_i(x, y \mid NDRS) = E_i(x, y \mid VRS)$;

(b) IRS $\Leftrightarrow E_i(x, y \mid NIRS) < E_i(x, y \mid NDRS) \leq E_i(x, y \mid VRS)$;

(c) DRS $\Leftrightarrow E_i(x, y \mid NDRS) < E_i(x, y|NIRS) \leq E_i(x, y \mid VRS)$;

(d) SCRS $\Leftrightarrow E_i(x, y \mid NIRS) = E_i(x, y \mid NDRS) < E_i(x, y \mid VRS)$;

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.

We now develop a new 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.[18]

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 (MRAC) with respect to technology T is given by

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

[18] Our approach is similar but more general than the one developed in Cesaroni and Giovannola (2015).

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

The following lemma provides some relations between MRAC 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) $\mathrm{MRAC}(y, w \mid \mathrm{CRS}) = C(y, w \mid \mathrm{CRS})$*;*

(b) $\mathrm{MRAC}(y, w \mid \mathrm{VRS}) = C(y, w \mid \mathrm{CRS})$*;*

(c) $\mathrm{MRAC}(y, w \mid \mathrm{CRS}) = \mathrm{MRAC}(y, w \mid \mathrm{VRS}) = C(y, w \mid \mathrm{CRS})$*;*

(d) $\mathrm{MRAC}(y, w \mid T) \leq C(y, w \mid T)$*.*

The proof of Lemma 3.2 is in the Appendix A. Using MRAC, 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 $C(y, w \mid \mathrm{VRS})$:

(a) CES if $\mathrm{MRAC}(y, w \mid \mathrm{VRS})$ is attained for a single optimal scale size $\lambda = 1$;

(b) IES if $\mathrm{MRAC}(y, w \mid \mathrm{VRS})$ is attained for a single optimal scale size $\lambda > 1$;

(c) DES if $\mathrm{MRAC}(y, w \mid \mathrm{VRS})$ is attained for a single optimal scale size $\lambda < 1$;

(d) SCES if $\mathrm{MRAC}(y, w \mid \mathrm{VRS})$ is attained for multiple optimal scale sizes, some with $\lambda > 1$ and others with $\lambda < 1$,

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

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 $C(y, w \mid \mathrm{VRS})$ hold true:*

(a) CES if $C(y, w \mid NIRS) = C(y, w \mid NDRS) = C(y, w \mid VRS)$;

(b) IES if $C(y, w \mid NIRS) < C(y, w \mid NDRS) \leq C(y, w \mid VRS)$;

(c) DES if $C(y, w \mid NDRS) < C(y, w \mid NIRS) \leq C(y, w \mid VRS)$*;*

(d) SCES if $C(y, w \mid NIRS) = C(y, w \mid NDRS) < C(y, w \mid VRS)$*.*

The proof of Proposition 3.2 is in the Appendix A. Observe that the structure is very similar to Proposition 3.1 dealing with returns to scale in production correspondences.

4 Description of the Samples

We employ two secondary data sets for our empirical analysis.[19] 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 inter-temporal 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).

As a second sample we draw upon an unbalanced panel of three years (1984-1986) of French fruit producers based on annual accounting data collected in a survey (Ivaldi, Ladoux, Ossard, and Simioni (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, Ladoux, Ossard, and Simioni (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.

[19] Data are taken from the Journal of Applied Econometrics Data archive (see `http://qed.econ.queensu.ca/jae/`).

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. The last two columns report the difference in terms of the nonconvex estimates (i.e., $(C^{NC}(y,w) - C^{C}(y,w))/C^{NC}(y,w))$. 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 on the VRS technology.

A detailed analysis of the last two columns can furthermore reveal us something on the impact of multiple outputs and constant returns to scale 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 constant returns to scale, 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 constant returns to scale 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.

Table 2 about here

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. Figures 1(a) and 1(b) represent the densities for most part of the observed range of cost estimates for the hydro-power plants and the fruit producers respectively.[20] Figure 1(a) illustrates clearly: (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 1(b) reveals downwardly asymmetrical densities for all frontiers and a markedly higher mode for the convex cases for a given returns to scale assumption.

Figure 1 about here

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 Output

As a graphical illustration of the above sample level results, we plot convex and nonconvex cost functions for some specific observations. Figure 2 plots cost estimated on VRS along the single output for hydro-power plant number 16 in July 1997. Figure 4 plots costs estimated on VRS along the first output dimension solely for fruit producer number 19 in 1984. Figures 3 and 5 plot total costs for the same individual observations for the CRS case for the respective

[20] To enhance comparability we use a common Sheather and Jones plug-in bandwidth for the convex and nonconvex data series to be compared (see, e.g., Sheather (2004)).

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 2, 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 Figure 4 (again apart from the extremes).[21] Second, for observations within the empirical range the nonconvex cost function always offers a substantially closer fit compared to its convex counterpart. Representing the original observation by a bullet ($\bullet$), the nonconvex and convex projection points on the cost frontiers are indicated by a times sign ($\times$) and circle ($\circ$) respectively. One can observe that the difference between cost frontiers is huge in Figure 2 for observation 16 in July 1997. This difference is slightly less pronounced for observation 19 in Figure 4.

For the CRS case, both projection points coincide in the single output case in Figure 3, but diverge in the multiple output case in Figure 5. 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 Figures 2 and 4).

One possible indication of the global fit is a count of the number of change points spanning the piecewise linear cost functions. For the VRS cost functions depicted in Figures 2 and 4, 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 Figures 3 and 5, one finds on the one hand 0 and 0, and on the other hand 0 and 16 change points for

[21] 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 precise structure of the sample. This intuition for the single output case is developed in the figures in Appendix B by computing the models in Section 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 C for figures computed on a numerical example with one input and two outputs).

the convex and nonconvex cases, respectively.[22] Clearly, the global fit of the nonconvex cost function is way better than that of its convex counterparts.

To escape the potentially anecdotal nature of this evidence, Table 3 reports descriptive statistics on the number of change points for all observations in both samples.[23] 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 Figures 2 to 5 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 maintained.

Table 3 about here

Figures 2 to 5 about here

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) and Park, Jeong, Shin, and Lee (2010)). 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.[24] 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.

[22] The numerical procedure to determine these change points starts from a refined grid of 10000 points.

[23] To reduce the computational burden, the numerical procedure to determine the change points starts from a grid of 1000 points only.

[24] 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.

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. Table 4 summarizes the results per sample for both a production- and cost-based analysis. First, the majority of hydro-power plant and fruit producer observations are subjected to increasing returns to scale. One qualification is that the nonconvex cost approach for the hydro-power plants in fact indicates about an equal amount of increasing returns to scale and decreasing returns to scale. For both the hydro-power plants and the fruit producers the nonconvex cost approach reveals a larger share of observations subject to decreasing returns to scale compared to the production-based analysis. Second, there are more observations with constant returns to scale 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.

Table 4 about here

A natural question to ask is to what extent these overall results hide any differences between convex and nonconvex approaches, which are the main focus of this study. Per data set and per production and cost method, we report in Table 5 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 5, 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 sum to the corresponding numbers in Table 4 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.

We make two basic observations on these results in Table 5. 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.

Table 5 about here

6 Conclusions

Apart from a new proposition to determine scale economies for cost frontiers evaluated relative to non-convex technologies, 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. 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. 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, Rossi, and Ruzzier (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)).[25] 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 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.

An open question is to what extent existing empirical methodologies need to be re-

[25]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.

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. The lack of standards to report nonconvex results may well contribute to its negligence.

References

AFRIAT, S. (1972): "Efficiency Estimation of Production Functions," *International Economic Review*, 13(3), 568–598.

ALAM, I., AND R. SICKLES (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., C. NG, T. CHENG, AND M. KOVALYOV (2008): "A Survey of Scheduling Problems with Setup Times or Costs," *European Journal of Operational Research*, 187(3), 985–1032.

ATKINSON, S., AND J. DORFMAN (2009): "Feasible Estimation of Firm-Specific Allocative Inefficiency Through Bayesian Numerical Methods," *Journal of Applied Econometrics*, 24(4), 675–697.

BADUNENKO, O., D. HENDERSON, AND S. KUMBHAKAR (2012): "When, Where and How to Perform Efficiency Estimation," *Journal of the Royal Statistical Society*, 175A(4), 863–892.

BALAGUER-COLL, M., D. PRIOR, AND E. TORTOSA-AUSINA (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., AND K. JÖRNSTEN (2008): "Equilibrium Prices Supported by Dual Price Functions in Markets with Non-Convexities," *European Journal of Operational Research*, 190(3), 768–789.

Briec, W., K. Kerstens, and P. Vanden Eeckaut (2004): "Non-convex Technologies and Cost Functions: Definitions, Duality and Nonparametric Tests of Convexity," *Journal of Economics*, 81(2), 155–192.

Budish, E., P. Cramton, and J. Shim (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., and D. Giovannola (2015): "Average-Cost Efficiency and Optimal Scale Sizes in Non-Parametric Analysis," *European Journal of Operational Research*, 242(1), 121–133.

Copeland, A., and G. Hall (2011): "The Response of Prices, Sales, and Output to Temporary Changes in Demand," *Journal of Applied Econometrics*, 26(2), 232–269.

Coviello, D., A. Ichino, and N. Persico (2014): "Time Allocation and Task Juggling," *American Economic Review*, 104(2), 609–623.

Cummins, D., and H. Zi (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., and K. Kerstens (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., and K. Fox (2014): "Reference Technology Sets, Free Disposal Hulls and Productivity Decompositions," *Economics Letters*, 122(2), 238–242.

Diewert, W., and C. Parkan (1983): "Linear Programming Test of Regularity Conditions for Production Functions," in *Quantitative Studies on Production and Prices*, ed. by W. Eichhorn, K. Neumann, and R. Shephard, pp. 131–158. Physica-Verlag, Würzburg.

Eaton, B., and R. Lipsey (1997): *On the Foundations of Monopolistic Competition and Economic Geography: The Selected Essays of B. Curtis Eaton and Richard G. Lipsey.* Edward Elgar, Cheltenham.

Estache, A., M. Rossi, and C. Ruzzier (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., AND A. ULLAH (1999): "On Goodness-of-fit Tests for Weakly Dependent Processes Using Kernel Method," *Journal of Nonparametric Statistics*, 11(1), 337–360.

FARRELL, M. (1959): "The Convexity Assumption in the Theory of Competitive Markets," *Journal of Political Economy*, 67(4), 377–391.

FIRST, Z., S. HACKMAN, AND U. PASSY (1993): "Efficiency Estimation and Duality Theory for Nonconvex Technologies," *Journal of Mathematical Economics*, 22(3), 295–307.

FUSS, M., D. MCFADDEN, AND Y. MUNDLAK (1978): "A Survey of Functional Forms in the Economic Analysis of Production," in *Production Economics: A Dual Approach to Theory and Applications, vol. 1*, ed. by M. Fuss, and D. McFadden, pp. 219–268. North-Holland, Amsterdam.

GARBACCIO, R., B. HERMALIN, AND N. WALLACE (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., AND K. KERSTENS (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 Revisted," *Quarterly Journal of Economics*, 93(1), 107–129.

HACKMAN, S. (2008): *Production Economics: Integrating the Microeconomic and Engineering Perspectives*. Springer, Berlin.

HACKMAN, S., U. PASSY, AND L. PLATZMAN (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., C. LE VAN, AND P. MICHEL (2009): "Non-convex aggregate technology and optimal economic growth," *Economic Theory*, 40(3), 457–471.

IVALDI, M., N. LADOUX, H. OSSARD, AND M. SIMIONI (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., G. JOHNES, R. OSKROCHI, AND R. CROUCHLEY (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., AND S. MANAGI (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., AND P. VANDEN EECKAUT (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., O. UTKIN, A. VOLODIN, I. SABLIN, AND M. PATRIN (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.

LI, Q. (1996): "Nonparametric Testing of Closeness between Two Unknown Distribution Functions," *Econometric Reviews*, 15(1), 261–274.

MADDALA, G., AND R. ROBERTS (1981): "Statistical Cost Analysis Re-Revisited: Comment," *Quarterly Journal of Economics*, 96(1), 177–182.

MAS-COLELL, A., A. WHINSTON, AND J. GREEN (1995): *Microeconomic Theory*. Oxford University Press, Oxford.

MÜSER, M., AND H. DYCKHOFF (2017): "Quality Splitting in Waste Incineration due to Non-convex Production Possibilities," *Journal of Business Economics*, 87(1), 73–96.

PARK, J.-B., Y.-W. JEONG, J.-R. SHIN, AND K. LEE (2010): "An Improved Particle Swarm Optimization for Nonconvex Economic Dispatch Problems," *IEEE Transactions on Power Systems*, 25(1), 156–166.

PODINOVSKI, V. (2004a): "Efficiency and Global Scale Characteristics on the "No Free Lunch" Assumption Only," *Journal of Productivity Analysis*, 22(3), 227–257.

——— (2004b): "Local and Global Returns to Scale in Performance Measurement," *Journal of the Operational Research Society*, 55(2), 170–178.

Preckel, P., and T. Hertel (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.

——— (2004): *Data Envelopment Analysis: Theory and Techniques for Economics and Operations Research*. Cambridge University Press, Cambridge.

Ray, S., and H. Kim (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., and K. Mukherjee (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 *Social Choice and Public Decision Making: Essays in Honor of Kenneth J. Arrow, Volume I*, ed. by W. Heller, R. Starr, and S. Starrett, pp. 117–134. Cambridge University Press, Cambridge.

——— (1994): "The Allocation of Resources in the Presence of Indivisibilities," *Journal of Economic Perspectives*, 8(4), 111–128.

Sheather, S. (2004): "Density Estimation," *Statistical Science*, 19(4), 588–597.

SHEPHARD, R. (1970): *Theory of Cost and Production Functions*. Princeton University Press, Princeton.

——— (1974): *Indirect Production Functions*. Verlag Anton Hain, Meisenheim am Glam.

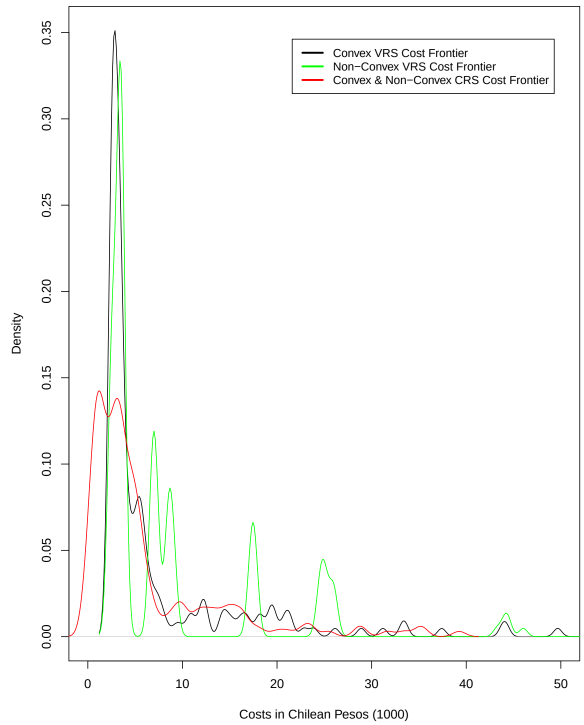
VARIAN, H. (1984): "The Nonparametric Approach to Production Analysis," *Econometrica*, 52(3), 579–597.

VITON, P. (2007): "Cost Efficiency in US Air Carrier Operations, 1970-1984: A Comparative Study," *International Journal of Transport Economics*, 34(3), 369–401.

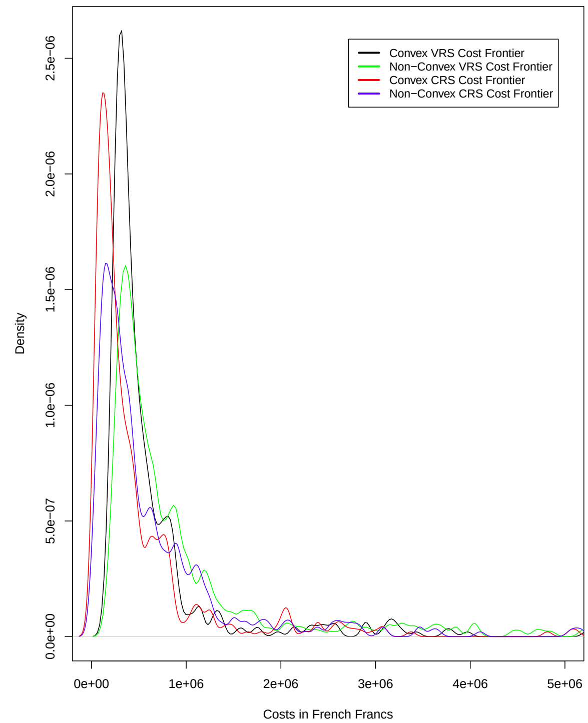
WIBE, S. (1984): "Engineering Production Functions: A Survey," *Economica*, 51(204), 401–411.

WINTER, S. (2008): "Scaling Heuristics Shape Technology! Should Economic Theory Take Notice?," *Industrial and Corporate Change*, 17(3), 513–531.

Figure 1: Kernel Density Estimates of Cost Function Estimates for (a) Chilean Hydro-power Plants and (b) French Fruit Producers



(a)



(b)

Figure 2: VRS Cost Function in the Single Output for Hydro-power Plant 16 in July 1997

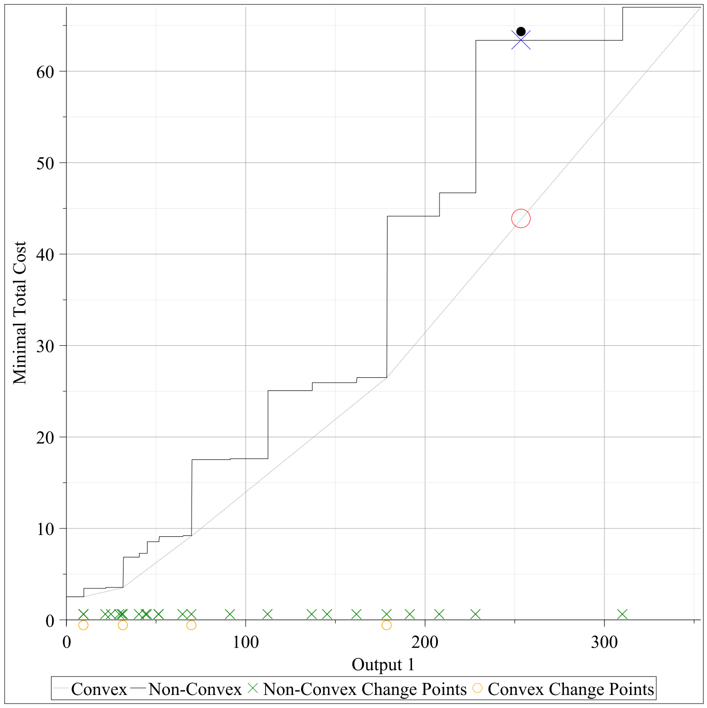


Figure 3: CRS Cost Function in the Single Output for Hydro-power Plant 16 in July 1997

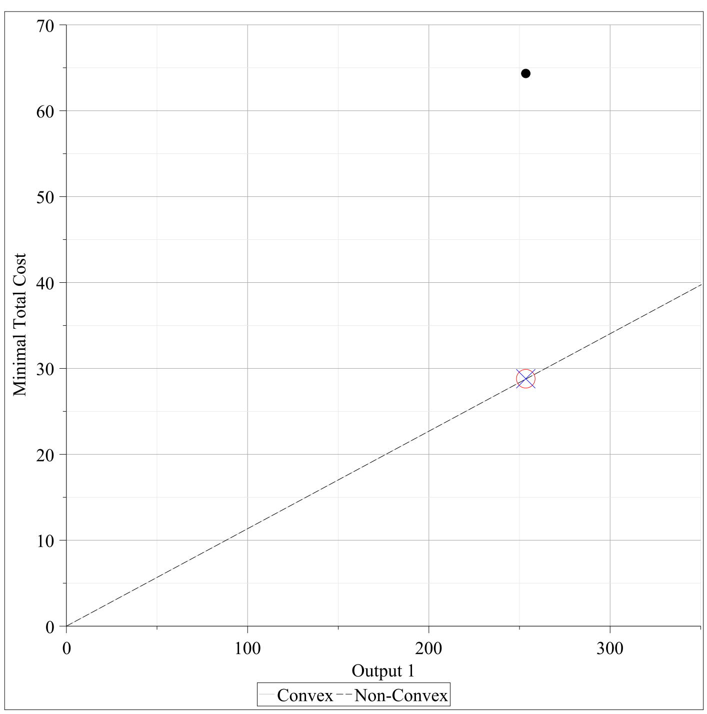


Figure 4: VRS Cost Function in Output 1 for Fruit Producer 19 in 1984

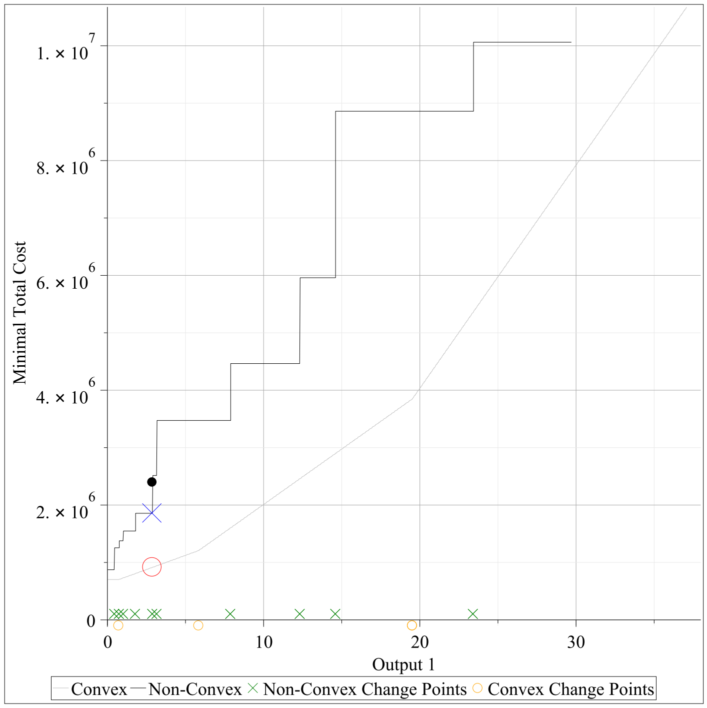
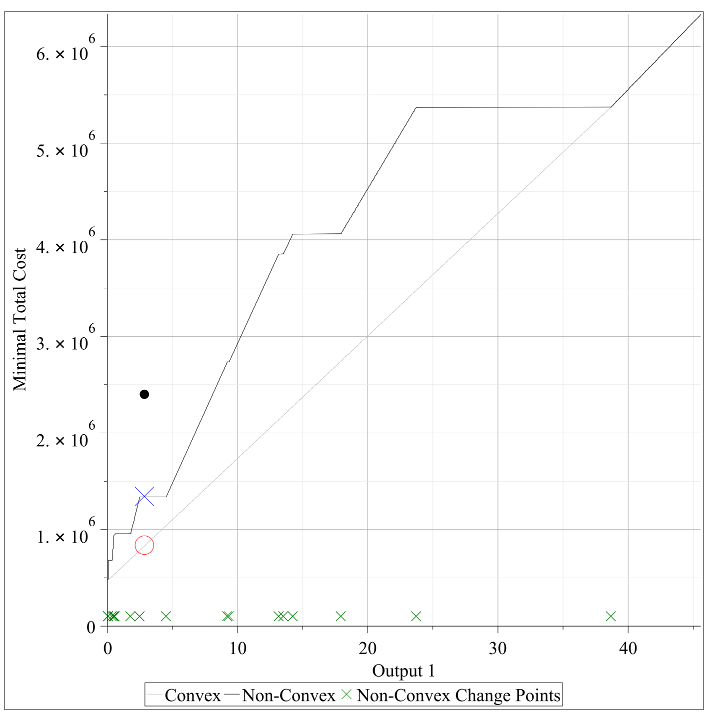


Figure 5: CRS Cost Function in Output 1 for Fruit Producer 19 in 1984



Article	Δ w.r.t. NC	Remarks
Nonparametric Frontier Studies:		
Balaguer-Coll et al. (2007)	41.14%	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 & Zi (1998)	49.45%	sample
	54.35%	1988
	49.45%	1989
	50.00%	1990
	47.83%	1991
	45.65%	1992
De Borger & Kerstens (1996)	22.41%	
Garbaccio et al. (1994)	23.46%	Solvent
	22.37%	Insolvent
Grifell-Tatjé & Kerstens (2008)	9.15%	Actual
	20.18%	Ideal
Ray & 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
Alternative Studies:		
Copeland & Hall (2011)	4.36%	Dynamic motor vehicle production model
Ray (1997)	24.73%	1 Output

† pop = municipality population.

Table 1: Nonconvex and Convex Cost Estimates: Literature Review

Sample 1: Chilian Hydro-power Plants						
Costs	Nonconvex (NC)		Convex (C)		Δ w.r.t. NC	
	VRS	*CRS*	*VRS*	*CRS*	*VRS*	*CRS*
Average	10.6663	6.2228	8.4522	6.2228	20.76 %	0.00 %
Stand.Dev.	13.6422	7.4285	11.0802	7.4285	18.78 %	0.00 %
Minimum	2.4912	0.0442	2.4912	0.0442	0.00 %	0.00 %
Maximum	65.8023	39.2475	65.8023	39.2475	0.00 %	0.00 %
Li test†	VRS 12.18***				CRS 0.0	
Sample 2: French Fruit Producers						
Costs	Nonconvex (NC)		Convex (C)		Δ w.r.t. NC	
	VRS	*CRS*	*VRS*	*CRS*	*VRS*	*CRS*
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	13448.4	6754.19	11815.7	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 (*).

Table 2: Nonconvex and Convex Cost Function Estimates: Descriptive Statistics

Sample 1: Chilian Hydro-power Plants				
Costs	Nonconvex (NC)		Convex (C)	
	VRS	*CRS*	*VRS*	*CRS*
Average	19.44	0	4.13	0
Stand.Dev.	0.71	0	0.34	0
Minimum	19	0	4	0
Maximum	21	0	5	0
Sample 2: French Fruit Producers				
Costs	Nonconvex (NC)		Convex (C)	
Ouput 1	*VRS*	*CRS*	*VRS*	*CRS*
Average	13.16	10.74	4.88	1.92
Stand.Dev.	2.94	2.94	0.95	0.28
Minimum	0	1	0	1
Maximum	18	19	8	2
Ouput 2	*VRS*	*CRS*	*VRS*	*CRS*
Average	5.95	9.75	3.69	0.98
Stand.Dev.	2.11	2.92	1.20	0.13
Minimum	0	1	0	0
Maximum	10	20	6	1

Table 3: Nonconvex and Convex Cost Function Sections: Descriptive Statistics on Change Points

Chilean Hydro-power Plants (%)			
Production	*IRS*	*CRS*	*DRS*
Nonconvex	70.31	16.67	13.02
Convex	76.04	2.6	21.35
Cost	*IES*	*CES*	*DES*
Nonconvex	51.56	0.52	47.92
Convex	68.23	9.38	22.4
French Fruit Producers (%)			
Production	*IRS*	*CRS*	*DRS*
Nonconvex	74.07	12.84	13.09
Convex	90.37	1.73	7.9
Cost	*IES*	*CES*	*DES*
Nonconvex	73.83	1.98	24.2
Convex	93.33	0.25	6.42

Table 4: Returns to Scale and Economies of Scale Results

Chilean Hydro-power Plants (%)				
Production	Convex			
Nonconvex	*IRS*	*CRS*	*DRS*	Total
IRS	63.54	0	6.77	70.31
CRS	11.46	2.6	2.6	16.67
DRS	1.04	0	11.98	13.02
Total	76.04	2.6	21.35	100
Agreement				78.13
Disagreement				21.88
Cost	Convex			
Nonconvex	*IES*	*CES*	*DES*	Total
IES	46.35	5.21	0	51.56
CES	0	0.52	0	0.52
DES	21.88	3.65	22.4	47.92
Total	68.23	9.38	22.4	100
Agreement				69.27
Disagreement				30.73
French Fruit Producers (%)				
Production	Convex			
Nonconvex	*IRS*	*CRS*	*DRS*	Total
IRS	74.07	0	0	74.07
CRS	9.38	1.48	1.98	12.84
DRS	6.91	0.25	5.93	13.09
Total	90.37	1.73	7.9	100
Agreement				81.48
Disagreement				18.52
Cost	Convex			
Nonconvex	*IES*	*CES*	*DES*	Total
IES	73.83	0	0	73.83
CES	1.73	0.25	0	1.98
DES	17.78	0	6.42	24.2
Total	93.33	0.25	6.42	100
Agreement				80.49
Disagreement				19.51

Table 5: Returns to Scale and Economies of Scale: Convex vs. Nonconvex Models

Appendices

A Proof of propositions

Proof of Lemma 3.2:

Proof. (a) Using Definition 3.1 and the cost function (6),

$$\begin{aligned}\mathrm{MRAC}(y, w \mid \mathrm{CRS}) &= \min_{\lambda}\left\{\frac{C(\lambda y, w \mid \mathrm{CRS})}{\lambda} \,:\, \lambda > 0\right\}\\ &= \min_{z,x,\delta,\lambda}\left\{\frac{wx}{\lambda} \,:\, x \geq \sum_{k=1}^{K} x_k \delta z_k, \lambda y \leq \sum_{k=1}^{K} y_k \delta z_k, z \in \Lambda, \delta \geq 0\right\}.\end{aligned}$$

Denote $x' = \frac{x}{\lambda}$ and $\delta' = \frac{\delta}{\lambda}$. Then the previous expression can be rewritten as

$$\begin{aligned}\mathrm{MRAC}(y, w \mid \mathrm{CRS}) &= \min_{z,x,\delta',\lambda}\left\{wx' \,:\, \lambda x' \geq \sum_{k=1}^{K} x_k \delta z_k, \lambda y \leq \sum_{k=1}^{K} y_k \delta z_k, z \in \Lambda, \delta \geq 0\right\}\\ &= \min_{z,x',\delta'}\left\{wx' \,:\, x' \geq \sum_{k=1}^{K} x_k \delta' z_k, y \leq \sum_{k=1}^{K} y_k \delta' z_k, z \in \Lambda, \delta' \geq 0\right\}\\ &= C(y, w \mid \mathrm{CRS}).\end{aligned}$$

(b) Using Definition 3.1 and the cost function (6),

$$\begin{aligned}\mathrm{MRAC}(y, w \mid \mathrm{VRS}) &= \min_{\lambda}\left\{\frac{C(\lambda y, w \mid \mathrm{VRS})}{\lambda} \,:\, \lambda > 0\right\}\\ &= \min_{z,x,\lambda}\left\{\frac{wx}{\lambda} \,:\, x \geq \sum_{k=1}^{K} x_k z_k, \lambda y \leq \sum_{k=1}^{K} y_k z_k, z \in \Lambda\right\}.\end{aligned}$$

Denote $x' = \frac{x}{\lambda}$ and $\delta' = \frac{1}{\lambda}$. Then the previous expression can be rewritten as

$$\begin{aligned}\mathrm{MRAC}(y, w \mid \mathrm{VRS}) &= \min_{z,x,\lambda}\left\{wx' \,:\, \lambda x' \geq \sum_{k=1}^{K} x_k z_k, \lambda y \leq \sum_{k=1}^{K} y_k z_k, z \in \Lambda\right\}\\ &= \min_{z,x',\delta'}\left\{wx' \,:\, x' \geq \sum_{k=1}^{K} x_k \delta' z_k, y \leq \sum_{k=1}^{K} y_k \delta' z_k, z \in \Lambda, \delta' \geq 0\right\}\\ &= C(y, w \mid \mathrm{CRS}).\end{aligned}$$

(c) Obviously, this case follows from combining (a) and (b).

(d) This results follows directly from Definition 3.1 by setting $\lambda = 1$. □

Proof of Proposition 3.2:

Proof.

(a) From Lemma 3.2(b), it follows that $\text{MRAC}(y, w \mid \text{VRS}) = C(y, w \mid \text{CRS})$. Using Lemma 3.1 and the assumption made in (a), we observe that $C(y, w \mid \text{CRS}) = \min\{C(y, w \mid \text{NIRS}), C(y, w \mid \text{NDRS})\} = C(y, w \mid \text{VRS})$. Combined, this results in $\text{MRAC}(y, w \mid \text{VRS}) = C(y, w \mid \text{VRS})$. Obviously, the latter is realized for the optimal scale size $\lambda = 1$ as can be observed from Definition 3.1. According to Definition 3.2, $C(y, w \mid \text{VRS})$ exhibits CES.

(b) From Lemma 3.2(b), it follows that $\text{MRAC}(y, w \mid \text{VRS}) = C(y, w \mid \text{CRS})$. Using Lemma 3.1 and the assumption made in (b), we observe that $C(y, w \mid \text{CRS}) = \min\{C(y, w \mid \text{NIRS}), C(y, w \mid \text{NDRS})\} = C(y, w \mid \text{NIRS})$. Consequently, $\text{MRAC}(y, w \mid \text{VRS}) = C(y, w \mid \text{NIRS})$. The latter can be rewritten using (6) as

$$\text{MRAC}(y, w \mid \text{VRS}) = \min_{z, x\delta} \left\{ wx \,:\, x \geq \sum_{k=1}^{K} x_k \delta z_k, y \leq \sum_{k=1}^{K} y_k \delta z_k, z \in \Lambda, 0 \leq \delta \leq 1 \right\}.$$

Denote $\lambda = \frac{1}{\delta}$ and $x' = \lambda x$. Then, the previous expression can be rewritten as

$$\begin{aligned}
\text{MRAC}(y, w \mid \text{VRS}) &= \min_{z, x, \lambda} \left\{ wx \,:\, \lambda x \geq \sum_{k=1}^{K} x_k z_k, \lambda y \leq \sum_{k=1}^{K} y_k z_k, z \in \Lambda, \lambda \geq 1 \right\} \\
&= \min_{z, x', \lambda} \left\{ \frac{wx'}{\lambda} \,:\, x' \geq \sum_{k=1}^{K} x_k z_k, \lambda y \leq \sum_{k=1}^{K} y_k z_k, z \in \Lambda, \lambda \geq 1 \right\} \\
&= \text{MRAC}(y, w \mid \text{NDRS}).
\end{aligned}$$

Obviously, the latter is realized for optimal scale size $\lambda \geq 1$. Excluding the case of $\lambda = 1$ which corresponds with CES, we observe IES according to Definition 3.2.

(c) The proof is analogous to that of (b) (switching NIRS with NDRS and changing the inequalities related to λ) and is left to the reader.

(d) The proof is a combination of the reasoning followed in (b) and (c), hereby excluding case (a) and is therefore left to the reader. □

B The case of 2 inputs and 1 output

Consider a technology with two inputs X_1 and X_2 and one output Y with 5 observations listed in Table A.1. The second and third columns represent the inputs X_1 and X_2, while the last column provides the corresponding output Y. This small numerical example serves as a basis to numerically compute and visualize the variable returns to scale (VRS) convex and nonconvex production and cost models in the main body of the text using Maple 18.

Observations	Input 1 (X_1)	Input 2 (X_2)	Output (Y)
Unit 1	1	1	1
Unit 2	1.5	1.2	2
Unit 3	1	2	3
Unit 4	0.5	2.5	1.4
Unit 5	3	3	3.5

Table A.1: Numerical Example with 2 Inputs and 1 Output

Obviously, the convex and nonconvex frontiers determined by this VRS-technology can be represented by 3-dimensional surfaces in the 3-dimensional inputs-output space. In Figure A.1a, an overlay of both these frontiers is shown with a slightly transparent convex frontier revealing the nonconvex frontier underneath. Cutting both these frontiers with a horizontal plane with equation $Y = 1.5$ in the same Figure A.1a gives rise to two planar curves, both of which can be observed in Figures A.1a and A.1b, the latter being the section in the two inputs space. These intersection curves (isoquants) with the convex and nonconvex frontiers are visible in Figure A.1b as dashed and solid curves respectively. Clearly, both curves do not have any points in common demonstrating that -not surprisingly- a convex technology need not necessarily approximate a nonconvex technology.

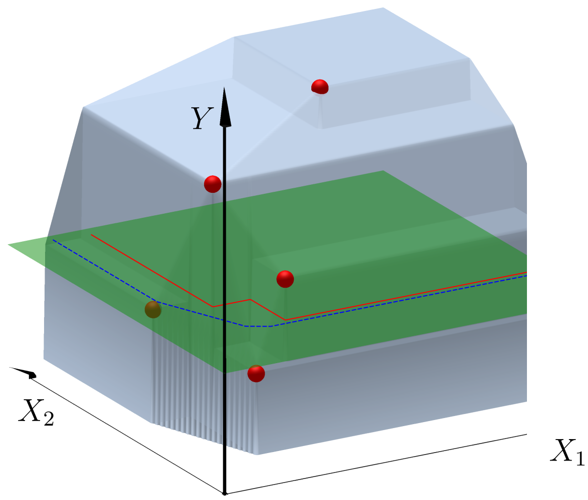
Let the unit input prices be fixed at 5 for both inputs allowing to compute cost functions. In Figure A.1b, these cost functions for convex and nonconvex technologies are visualized in the 2-dimensional input space as dashed and solid straight lines respectively. Obviously, both lines are parallel but do not coincide for this Y-level since the input combinations realizing the minimal cost (marked with circles) differ and are situated at two different isoquants.

Another way of representing these cost functions is by considering them as real-valued functions of the output Y. This yields the 2-dimensional representations provided in Figure

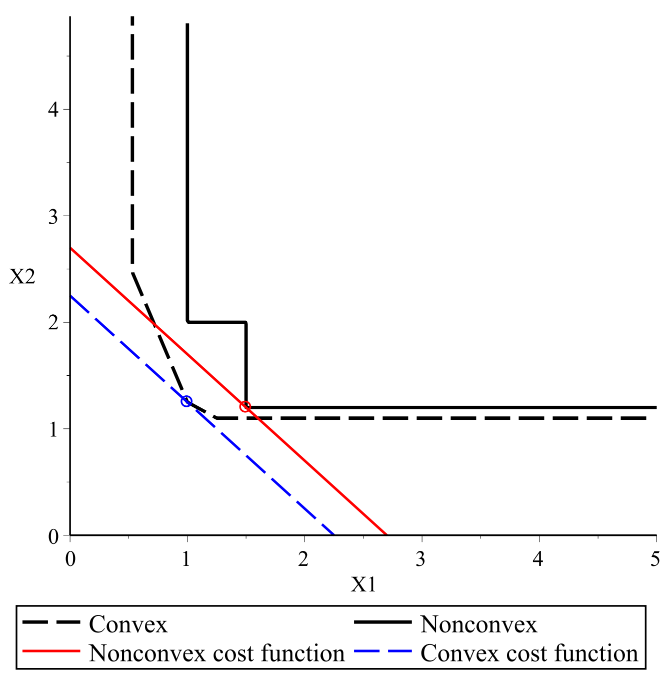
A.2a. Both Figure A.1b and A.2a confirm graphically the property that costs evaluated on convex technologies are smaller than or equal to costs evaluated on nonconvex technologies. Figure A.2a also shows that in the single output case convex and nonconvex cost functions share some common points. When looking at the input-output space, the path of all cost-minimizing combinations (X_1, X_2, Y) can be easily traced. These paths are shown in Figure A.2b as dashed and solid curves for convex and nonconvex technologies respectively.

[Figures A.1 and A.2 about here]

Figure A.1: Convex and nonconvex frontiers in input-output space for a VRS-technology with 2 inputs (X_1 and X_2) and 1 output (Y).

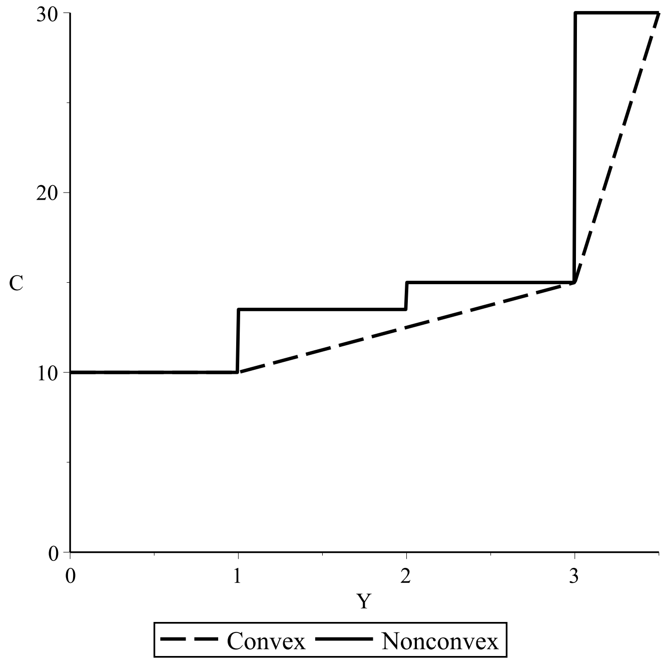


(a) 3-dimensional view with horizontal section at level $Y = 1.5$.

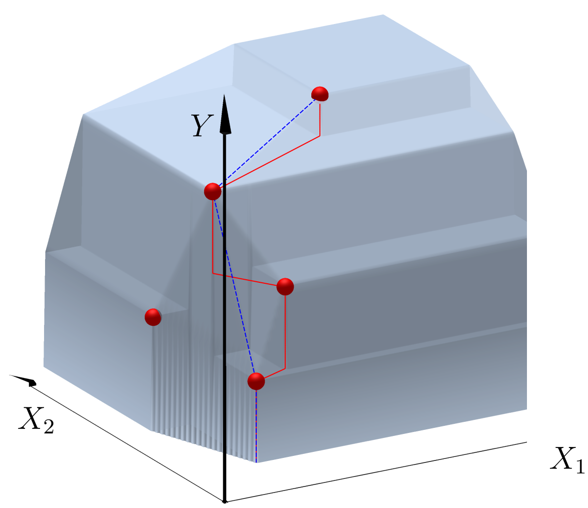


(b) 2-dimensional view of the section at level $Y = 1.5$ with corresponding cost functions.

Figure A.2: Convex and nonconvex cost functions for a VRS-technology with 2 inputs (X_1 and X_2) and 1 output (Y).



(a) Convex and nonconvex minimal cost functions in the single output.



(b) 3-dimensional view of the convex and nonconvex frontiers with paths of cost-minimizing input combinations.

C The case of 2 outputs and 1 input

Now we consider a technology with one input X and two outputs Y_1 and Y_2 with 5 observations listed in Table A.2. The second column refers to the input X, while the third and last columns reveal the corresponding outputs Y_1 and Y_2.

Observations	Input (X)	Output 1 (Y_1)	Output 2 (Y_2)
Unit 1	2	2	3
Unit 2	1.5	2.4	2
Unit 3	1	1	1.5
Unit 4	4	4.5	2.5
Unit 5	3	4	3

Table A.2: Numerical Example with 1 Input and 2 Outputs

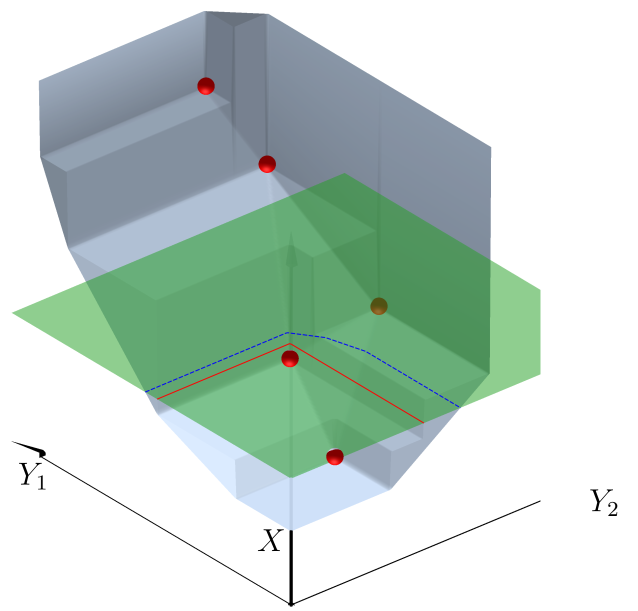
Also in this case, the convex and nonconvex frontiers determined by a variable returns to scale (VRS) technology can be represented by 3-dimensional surfaces in the 3-dimensional input-outputs space. In Figure A.3a, an overlay of both these frontiers is shown with a slightly transparent convex frontier revealing the nonconvex frontier underneath. Note that the axes determined by both outputs Y_1 and Y_2 are drawn in the horizontal plane and that the vertical axis represents the input X. Cutting these frontiers with a horizontal plane with equation $X = 1.7$ gives rise to two planar curves, both of which can be observed in Figure A.3a and A.3b. These output transformation curves with the convex and nonconvex frontiers are visible in Figure A.3b as dashed and solid curves respectively. Clearly, both output transformation curves do not have any point in common.

To compute cost functions, let the unit input price be fixed at 5. Then, the resulting cost functions can be seen as real-valued functions of the output combinations (Y_1, Y_2) resulting in a surface in (Y_1, Y_2, C)-space when visualized. Because of the single input in this numerical example, $C = 5X$ which implies that the graph of the cost functions is identical to the one of the VRS production frontier up to a scaling factor of 5 on the vertical axes. Considering this scaling factor, Figure A.3a can serve equally well as a visualization of the cost function in both outputs. Consequently, Figure A.3b serves as visualization of the iso-cost curves for convex (dashed curve) and nonconvex (solid curve) combinations of outputs for a fixed cost level of $C = 5 \times 1.7 = 8.5$.

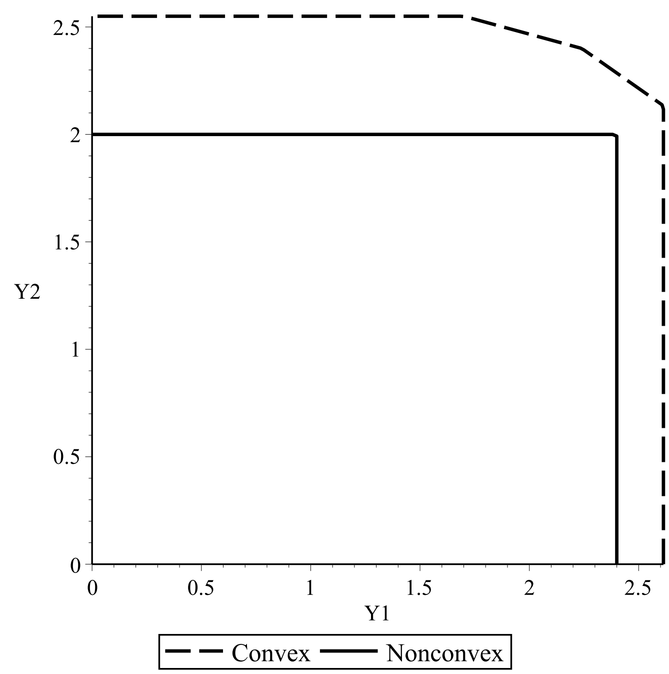
In Figure A.4a, the cost function is cut by a vertical plane with equation $Y_2 = 2.2$. This yields the 2-dimensional representations provided in Figure A.4b. In both figures, the sections with the convex and nonconvex cost function are shown as a dashed and solid curve, respectively. Clearly, it is confirmed that costs evaluated on convex technologies are smaller than or equal to costs evaluated on nonconvex technologies. Furthermore, Figure A.4b also shows that convex and nonconvex cost functions hardly share any common points except near one end of the empirical range of the data, which is in contrast with the single output case in Figure A.2a where points are in common at both ends of the empirical range and in between. Of course, in both the single and multi-output case, the exact location of common points seems mainly determined by the empirical configuration of the sample.

[Figures A.3 and A.4 about here]

Figure A.3: Convex and nonconvex frontiers in input-output space for a VRS-technology with 2 outputs (Y_1 and Y_2) and 1 input (X).

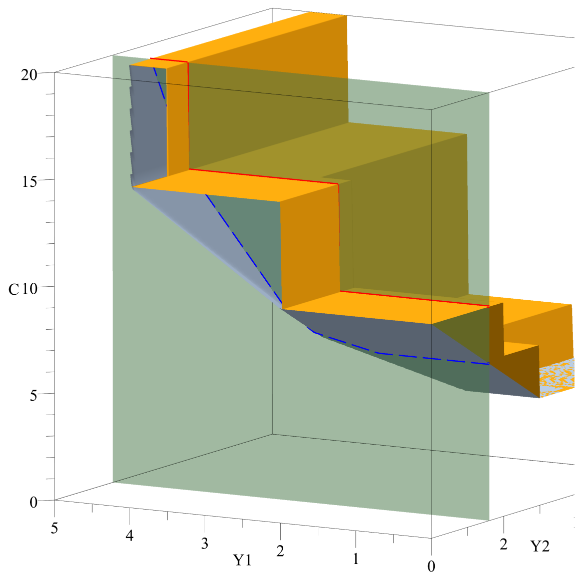


(a) 3-dimensional view with horizontal section at level $X = 1.7$.

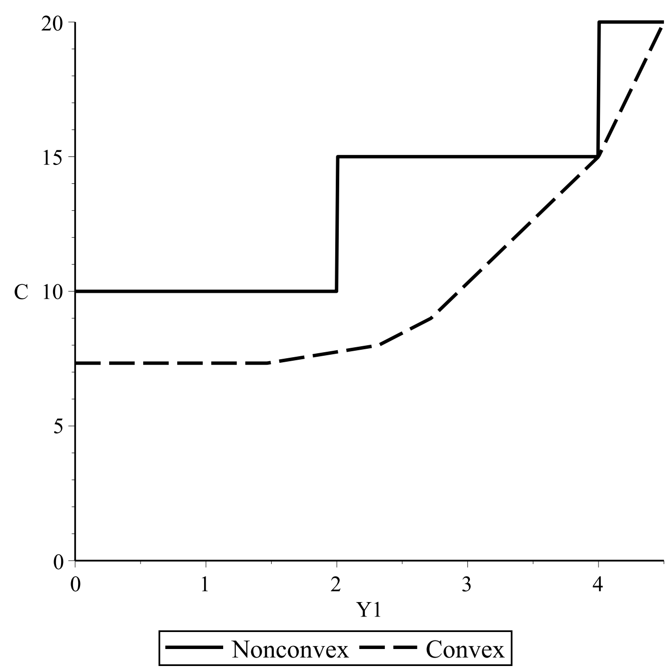


(b) 2-dimensional view of the section at level $X = 1.7$.

Figure A.4: Convex and nonconvex cost functions for a VRS-technology with 2 outputs (Y_1 and Y_2) and 1 input (X).



(a) 3-dimensional view with horizontal section at level $Y_2 = 2.2$.



(b) 2-dimensional view of the section at level $Y_2 = 2.2$.