Document de travail du LEM / Discussion paper LEM
2017- 18

# Convex and Nonconvex Input-Oriented Technical and Economic Capacity Measures: An Empirical Comparison

**Kristiaan KERSTENS**
LEM UMR 9221 / k.kerstens@ieseg.fr

**Jafar SADEGHI**
Department of Mathematics, Faculty of Mathematical Sciences and Computer, Kharazmi University

**Ignace VAN DE WOESTYNE**
KU Leuven, Research Unit MEES, Brussels, Belgium

http://lem.cnrs.fr/IMG/pdf/dp2017-18.pdf

http://lem.cnrs.fr/

# Convex and Nonconvex Input-Oriented Technical and Economic Capacity Measures: An Empirical Comparison

Kristiaan Kerstens*, Jafar Sadeghi†, Ignace Van de Woestyne‡

November 2017

**Abstract**

This contribution has two main objectives. First, it aims to compare empirically input-oriented technical and economic capacity notions. Second, it aims to compare these technical and economic capacity notions on both convex and nonconvex technologies. After defining these input-oriented technical and economic capacity notions, this contribution focuses on empirically comparing these different capacity utilization notions using a secondary data set. Anticipating two key empirical conclusions, we find that all these different capacity notions follow different distributions, and also that these distributions almost always differ under convex and nonconvex technologies.

**DRAFT: DO NOT QUOTE OR CITE WITHOUT PERMISSION**.

JEL codes: D24

Keywords: Data Envelopment Analysis; Technology; Cost function; Capacity utilization

*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

†Department of Mathematics, Faculty of Mathematical Sciences and Computer, Kharazmi University, 50, Taleghani Ave., 15618, Tehran, Iran, std_j.sadeghi@khu.ac.ir and visitor, CNRS-LEM (UMR 9221), IESEG School of Management, 3 rue de la Digue, F-59000 Lille, France

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

# 1 Introduction

Analysing efficiency and productivity using frontier technologies has become a standard empirical tool serving a variety of academic, regulatory and managerial purposes. Indeed there is a huge academic literature applying these methodologies for analyzing private and public sector performance-related issues. Focusing on empirical surveys of certain well-studied sectors, one can point, for example, to banking (Harker and Zenios (2001)), education (Worthington (2001)), health care (Ozcan (2008)), insurance (Cummins and Weiss (2000)), justice system (Voigt (2016)) and real estate (Anderson, Lewis, and Springer (2000)). Apart from this surge of empirical applications, there has equally been an extended series of methodological innovations in this literature surveyed in, for example, Hatami-Marbini, Emrouznejad, and Tavana (2011) or Thanassoulis, Silva Portela, and Despić (2008).

An important area of regulatory applications has been the implementation of incentive regulatory mechanisms (e.g., price cap regulation) using frontier-based performance benchmarks in countries with liberalized network industries (e.g., electricity, gas, water utilities). One survey focusing on its use in the electricity sector is Jamasb and Pollitt (2000). An example of a managerial application is the use of frontier methods to save resources allowing to use these as internal funds to pursue a growth strategy in a US bank (see, e.g., Sherman and Ladino (1995)).

However, this frontier literature has largely ignored integrating the important notion of capacity utilization. Consequently, part of what appears like inefficiency may in fact be due to the short-run fixity of certain inputs, depending on the exact definition of capacity utilization. It is of equal importance to account for heterogeneity in capacity utilization when measuring productivity growth (e.g., Luh and Stefanou (1991)).

Capacity utilization of fixed inputs is relevant for both managers and policy makers at various levels of aggregation and in all economic sectors. For instance, at the country level capacity utilization is traditionally employed as a leading macro-economic indicator to forecast inflation (e.g., Christiano (1981)). The management of excess vessel capacities has recently become a key policy issue in fisheries due to degrading bio-stocks in this common pool resource. As an example, a variety of capacity measures has been employed to evaluate vessel decommissioning schemes (e.g., Walden, Kirkley, and Kitts (2003)). To curb overfishing, governments must determine sustainable capacity levels by implementing a variety of policy measures (e.g., licenses, fishing day restrictions, etc.). To define these policy measures, scientists have developed short-run industry models based on vessel capacity estimates to allow planning the industry and infer realistic decommissioning schemes (see, e.g., Lindebo (2005)).

However, different notions of capacity co-exist in the literature (e.g., Christiano (1981) or Johansen (1968)). It is common to distinguish between technical or engineering concepts on the one hand and economic capacity concepts on the other hand. Johansen (1968) developed a technical or engineering approach by introducing a plant capacity notion. Plant capacity is defined as the maximal amount that can be produced per unit of time with existing plants and equipment without restrictions on the available variable inputs. This definition has been transposed into a production frontier context using output-oriented efficiency measures by Färe, Grosskopf, and Kokkelenberg (1989).

Most economic capacity concepts are based on the cost function. In the literature there are basically at least three ways of defining a cost-based capacity notion (see, e.g., Nelson (1989)). Each of these notions attempts to isolate the short-run inadequate or excessive utilization of fixed inputs. A first notion of potential outputs is defined in terms of the outputs produced at short-run minimum average total cost given existing plant and input prices (for instance, Hickman (1964)). It stresses the need to exploit scale economies in the short-run. A second definition of potential outputs is conceived in terms of the outputs produced at minimum average total cost in the long-run (e.g., Cassels (1937), among others). It is rarely used because its intertwining with the notion of scale economies. A third definition corresponds to the outputs at which the short-run and long-run average total cost curves are tangent. Since this tangency point is at the intersection of short-run and long-run expansion paths, this notion has considerable theoretical appeal (for example, Klein (1960) or Segerson and Squires (1990)).

We are unaware of any study comparing these technical and economic capacity notions. One plausible hypothesis explaining this lack of comparative studies is that the economic capacity notions at least implicitly adopt an input orientation, while the technical plant capacity notion is traditionally based on output-oriented efficiency measures. However, recently Cesaroni, Kerstens, and Van de Woestyne (2017) develop an input-oriented plant capacity notion based on input-oriented efficiency measures. Therefore, a first major goal of this contribution is to make a theoretically coherent input-oriented comparison between these technical and economic capacity notions. As a point of comparison, we also include the output-oriented plant capacity notion which has been used quite often in the literature in the last three decades since its inception (see Cesaroni, Kerstens, and Van de Woestyne (2017) for a literature review).

It is well-known that the axiom of convexity has a potential impact on the empirical analysis based on technologies (see, e.g., Tone and Sahoo (2003)). In our context, for instance, Walden and Tomberlin (2010) document the effect of maintaining or dropping convexity on the output-oriented plant capacity utilization concept. Equally so, Cesaroni, Kerstens, and Van de Woestyne (2017) reveal the impact of convexity on the input-oriented plant capacity utilization notion.

However, most researchers tend to ignore the potentially important impact of convexity on the cost function. This is related to a property of the cost function in the outputs that is ignored by most people. Indeed, some seminal contributions to axiomatic production theory indicate that the cost function is nondecreasing and convex in the outputs if and only if the technology is convex (e.g., Jacobsen (1970)). Otherwise, the cost function is nonconvex in the outputs. Briec, Kerstens, and Vanden Eeckaut (2004) refine this general property and prove that cost functions estimated on nonconvex technologies yield larger or equal cost estimates compared to cost functions estimated on convex technologies. These both types of cost functions are identical when there is a single output and constant returns to scale prevail. The large majority of empirical studies have ignored to put these properties to a test. In our context, to the best of our knowledge the impact of convexity on cost-based notions of capacity utilisation has never been evaluated. Therefore, a second major goal of this contribution is to make a coherent input-oriented comparison between technical and economic capacity notions using both convex and nonconvex technologies to assess the impact of the convexity hypothesis.

This contribution is structured as follows. Section 2 summarizes the basic definitions of the technology and the cost function. The next Section 3 reviews in detail both the economic and technical capacity utilization definitions. This includes, among others, looking at the issue of normalization, given the existence of inefficiencies, and a priori determining the eventual impact of convexity. In the next Section 4 we develop an empirical illustration making use of an existing secondary data set, which makes our results replicable. The focus is on descriptive statistics, a formal testing of the resulting distributions, and a comparison of Spearman rank correlations. A final section concludes.

## 2 Technology and Cost Functions: Basic Definitions

In this section we define technology and some basic notation. Given 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 $T$ can be defined as follows: $T = \{(x, y) | x \text{ can produce } y\}$. The input set associated with $T$ denotes all input vectors $x$ capable of producing a given output vector $y$: $L(y) = \{x | (x, y) \in T\}$. In a similar way, the output set associated with $T$ denotes all output vectors $y$ that can be produced from a given input vector $x$: $P(x) = \{y | (x, y) \in T\}$.

Throughout this contribution, technology $T$ satisfies some combination of the following standard assumptions:

(T.1) Possibility of inaction and no free lunch, i.e., $(0, 0) \in T$ and if $(0, y) \in T$, then $y = 0$.

(T.2) $T$ is a closed subset of $\mathbb{R}^N_+ \times \mathbb{R}^M_+$.

(T.3) Strong input and output disposal, i.e., 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) $\Gamma \equiv \Gamma^{\text{VRS}} = \{\delta : \delta = 1\}$.

(T.5) $T$ is convex.

Briefly discussing these traditional axioms on technology, it is useful to recall: (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) and variable returns to scale (VRS)), and (v) convexity of technology (see, e.g., Hackman (2008) for details). Not all these axioms are maintained in the empirical analysis.[1] In particular, key assumptions distinguishing some of the technologies in the empirical analysis are CRS versus VRS, and convexity versus nonconvexity.

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

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

The main properties of this input distance function 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)).

[1] Note that the convex VRS technology does not satisfy inaction.

The inverse of this input distance function $DF_i(x, y) = [D_i(x, y \mid T)]^{-1}$ is known as the radial input efficiency measure. Hence, the radial input efficiency measure is defined as:

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

Its key property is that it is situated between zero and unity ($0 < DF_i(x, y) \leq 1$), with efficient production on the boundary (isoquant) of the input set $L(y)$ represented by unity.

Switching to a dual representation of the technology, the cost function can be defined as the minimum expenditures needed to produce a given output vector $y$ for a given a vector of semi-positive input prices ($w \in \mathbb{R}^N_+$):

$$C(y, w \mid T) = \min_x\{wx : (x, y) \in T\} = \min_x\{wx : x \in L(y)\}. \tag{3}$$

Duality relations link these primal and dual representations 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 (3) is:

$$D_i(x, y \mid T) = \min_w\{wx : C(y, w \mid T) \geq 1\}, x \in L(y), \tag{4}$$

$$C(y, w \mid T) = \min_x\{wx : D_i(x, y \mid T) \geq 1\}, w > 0. \tag{5}$$

It is common to establish such duality relations under the hypothesis of a convex technology or a convex input set (e.g., (Hackman, 2008, Ch. 7)). Briec, Kerstens, and Vanden Eeckaut (2004) are the first to establish a local duality result between nonconvex technologies subject to various scaling laws and their corresponding nonconvex cost functions.

Next, the radial output efficiency measure can be defined as:

$$DF_o(x, y) = \max\{\theta \mid \theta \geq 0, \theta y \in P(x)\}. \tag{6}$$

It offers a complete characterization of the output set $P(x)$. Its main properties are that it is larger than or equal to unity ($DF_o(x, y) \geq 1$), with efficient production on the boundary (isoquant) of the output set $P(x)$ represented by unity, and that the radial output efficiency measure has a revenue interpretation (e.g., Hackman (2008)).

Partitioning the input vector into a fixed and variable part, we have ($x = (x^f, x^v)$) with $x^f \in \mathbb{R}^{N_f}_+$ and $x^v \in \mathbb{R}^{N_v}_+$ such that $N = N_f + N_v$. Furthermore, we can make the same distinction regarding the input price vector ($w = (w^f, w^v)$).

In a similar way, a short-run technology $T^f = \{(x^f, y) \in \mathbb{R}_+^{N_f} \times \mathbb{R}_+^M, x^f \text{ can produce } y\}$ and the corresponding input set $L^f(y) = \{x^f \in \mathbb{R}_+^{N_f} : (x^f, y) \in T^f\}$ and output set $P^f(x^f) = \{y | (x^f, y) \in T^f\}$ can be defined. Note that technology $T^f$ is obtained by a projection of technology $T \in \mathbb{R}_+^{N+M}$ into the subspace $\mathbb{R}_+^{N_f+M}$ (i.e., by setting all variable inputs equal to zero). The same applies by analogy to the input set $L^f(y)$ and the output set $P^f(x^f)$.

By analogy, the short-run total cost function is defined as follows:

$$C(w, x^f, y \mid T) = \min_{x^v} \{w^v x^v + w^f x^f : (x^v, x^f, y) \in T\}. \tag{7}$$

The short-run total cost function is simply the short-run variable cost function and the observed fixed costs. The short-run variable cost function is defined:

$$VC(w^v, x^f, y \mid T) = \min_{x^v} \{w^v x^v : (x^v, x^f, y) \in T\}. \tag{8}$$

The sub-vector input efficiency measure reducing only the variable inputs is defined as follows.

$$DF_i^{SR}(x^f, x^v, y) = \min\{\lambda \mid \lambda \geq 0, (x^f, \lambda x^v) \in L(y)\}. \tag{9}$$

Next, we need the following particular definitions of technologies $L(0) = \{x \mid (x, 0) \in S\}$ is the input set with zero output level. The sub-vector input efficiency measure reducing variable inputs evaluated relative to this input set with a zero output level is as follows.

$$DF_i^{SR}(x^f, x^v, 0) = \min\{\lambda \mid \lambda \geq 0, (x^f, \lambda x^v) \in L(0)\}. \tag{10}$$

Now, for $K$ observations $(x_k, y_k) \in \mathbb{R}_+^{N+M}, (k = 1, \ldots, K)$ a unified algebraic representation of convex and nonconvex nonparametric frontier technologies under CRS and VRS assumptions is possible as follows:

$$T^{\Lambda,\Gamma} = \left\{ (x, y) \,:\, x \geq \sum_{k=1}^{K} x_k \delta z_k, y \leq \sum_{k=1}^{K} y_k \delta z_k, z \in \Lambda, \delta \in \Gamma \right\}, \tag{11}$$

where

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

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

and

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

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

Observe there is one activity vector $z$ operating subject to a nonconvexity or convexity constraint as well as a scaling parameter $\delta$ allowing for some particular scaling of all $K$ observations determining the technology. The activity vector $z$ of real numbers summing to unity represents the convexity axiom, while this same sum constraint with each vector element being a binary integer is representing nonconvexity. The scaling parameter $\delta$ is free under CRS and fixed at the unit level under VRS.

To compute the input efficiency measure (2) or cost function (3) relative to convex technologies in (11) 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 literature (see Hackman (2008)).[2] For the nonconvex technologies, nonlinear binary mixed integer programs must be solved, but alternative solution strategies are available (see Kerstens and Van de Woestyne (2014)).

It is now useful to condition the above notation of the efficiency measures and cost functions relative to these nonparametric frontier technologies by distinguishing between constant (convention $CRS$) and variable (convention $VRS$) returns to scale assumptions, and between convexity (convention $C$) and nonconvexity (convention $NC$).

# 3 Economic and Technical Capacity Utilization: Literature Review and Definitions

A variety of capacity notions coexist in the economic literature. It is customary to distinguish between technical (engineering) and economic (mainly cost-based) capacity concepts (see, e.g., Johansen (1968); Nelson (1989)). We first address the economic concepts using a cost function approach, and then turn to the technical or engineering notion.

[2] By substituting $t_k = \delta z_k$ in (11), one can rewrite the sum constraint on the activity vector $z$. One must realize that the constraints on the scaling factor are integrated into the latter sum constraint and the LP appears.

## 3.1 Economic Capacity Concepts

At least three ways of defining a cost-based notion of capacity have been proposed in the literature (see Nelson (1989)). Each of these notions aims to isolate the short-run excessive or inadequate utilization of existing fixed inputs (e.g., capital stock). A first notion is defined in terms of the output produced at short-run minimum average total cost given existing input prices (see Hickman (1964), among others). A second definition focuses on the outputs for which short-run and long-run average total costs curves are tangent (e.g., Segerson and Squires (1990)). This tangency point notion is known under two variations depending on what are supposed to be the decision variables. One notion assumes that outputs are constant and determines optimal variable and fixed inputs. Another notion assumes that fixed inputs cannot adjust, but outputs, output prices and fixed input prices do adjust. A third and final definition of economic capacity considers the output determined by the minimum of the long-run average total costs (e.g., Cassels (1937), Klein (1960)).

To apply these notions of economic capacity utilization using nonparametric frontier technologies, one can characterize the above three economic capacity notions, one of which has two variants, in a multiple output context in the following series of definitions (see, e.g, De Borger, Kerstens, Prior, and Van de Woestyne (2012)).

**Definition 3.1.** The minimum of the short-run total cost function $C(y, w^v, x^f|VRS)$ (7) is $C(y, w^v, x^f|CRS)$.

The minimum of the single output short-run average total cost function can be determined indirectly in the multiple output case by solving for a variable cost function relative to a CRS technology ($VC(y, w^v, x^f|CRS)$), and simply adding observed fixed costs $FC = w^f x^f$. The resulting short-run total cost function $C(y, w^v, x^f|CRS)(= VC(y, w^v, x^f|CRS) + FC)$ offers the reference point for this capacity notion. In the convex case, computing a cost function boils down to a well-known linear program. But, in the nonconvex case one must solve a mixed binary integer linear program.

**Definition 3.2.** (i): Tangency cost with modified fixed inputs $C^{tang1}(y, w, x^{f*}|VRS)$ is $C(y, w|VRS) = C(y, w^v, x^{f*}|VRS)$.

(ii): Tangency cost with modified outputs $C^{tang2}(y(p, w^f, x^f), w, x^f|VRS)$ is $C(y(p, w^f, x^f), w|VRS) = C(y(p, w^f, x^f), w^v, x^f|VRS)$,

where $x^{f*}$ represents optimal fixed inputs, $p \in \mathbb{R}^N_+$ is a vector of input prices, and $y(p, w^f, x^f)$

represents outputs that have been adjusted in terms of given output prices, fixed input prices and the given fixed inputs.

First, the tangency point between short- and long-run costs can also be estimated using nonparametric cost frontiers. One can actually envision two types of tangency points depending on which variables one assumes to be decision variables.

One tangency cost notion assumes that outputs remain constant and then determines optimal variable and fixed inputs $C^{tang1}(y, w, x^{f*} \mid VRS)$. This can be solved indirectly by minimizing a long-run total cost function $C(y, w \mid VRS)$ yielding optimal fixed inputs $(x^{f*})$. By definition, the short-run and total cost function with fixed inputs equal to these ex post optimal fixed inputs $FC(y, w^v, x^{f*} \mid VRS)$ yields exactly the same solution in terms of optimal costs and optimal variable inputs $C(y, w^v, x^{f*} \mid VRS) = VC(y, w^v, x^{f*} \mid VRS) + FC(y, w^v, x^{f*} \mid VRS)$. Hence, the optimal solution for $C(y, w \mid VRS)$ generates the tangency point we are looking for. In the convex case, computing this cost function requires solving again a linear program. In the nonconvex case, one needs to solve a mixed binary integer linear programming problem.

Another tangency point, favored by Nelson (1989, p. 277) and analyzed in detail in Briec, Kerstens, Prior, and Van de Woestyne (2010), assumes that fixed inputs cannot be adjusted in the short-run, but that outputs, output prices ($p \in \mathbb{R}^M_+$) and fixed input prices are adjustable such that installed capacity is utilized ex post at a tangency cost level $(C^{tang2}(y(p, w^f, x^f), w, x^f \mid VRS))$. Though one may object that outputs are assumed to be exogenous in a competitive cost minimization model, this tangency notion offers a useful reference point, since it retrospectively indicates the output quantities and prices as well as the fixed input prices at which existing fixed inputs would have been optimally utilized. For an arbitrary observation, this tangency cost level may imply an output level $(y(p, w^f, x^f))$ below or above current outputs. In the convex case, optimal costs at this tangency point are determined by solving for each observation a nonlinear system of inequalities (Briec, Kerstens, Prior, and Van de Woestyne (2010)). In the nonconvex case, however, one must solve for each observation a mixed binary integer nonlinear system of inequalities.

**Definition 3.3.** The minimum of the long-run total cost function $C(y, w^v, x^f|VRS)$ is obtained as $C(y, w^v, x^f|CRS)$.

The minimum of long-run average total costs can be determined indirectly by solving for a long-run total cost function defined relative to a $CRS$ technology $C(y, w \mid CRS)$. In the convex case, computing this cost function again involves solving a linear program. For the nonconvex case, one must solve a mixed binary integer linear programming problem. For convenience, the way of computing all plant and economic capacity concepts in the convex as well as nonconvex case are spelled out in the Appendix.

In a frontier context, some of the above cost-based capacity concepts or some combination there-off have been reported in Giménez and Prior (2007), Prior-Jiménez (2003), or Sahoo and Tone (2009), among others. Note that we have ignored the discussion of alternative capacity concepts based on the revenue function (e.g., Lindebo, Hoff, and Vestergaard (2007)) or the profit function (e.g., Coelli, Grifell-Tatjé, and Perelman (2002)).

## 3.2 Plant Capacity Concepts

Johansen (1968) proposed a plant capacity notion that has been made operational by Färe, Grosskopf, and Kokkelenberg (1989) and Färe, Grosskopf, and Valdmanis (1989) using a pair of output-oriented efficiency measures. The plant capacity notion is defined by Johansen as "the maximum amount that can be produced per unit of time with existing plant and equipment, provided that the availability of variable factors of production is not restricted." Cesaroni, Kerstens, and Van de Woestyne (2017) develop a plant capacity notion using a pair of input-oriented efficiency measures. All of these proposals use $VRS$ technologies.

We now recall the definitions of the output- and input-oriented plant capacity utilization ($PCU$).

**Definition 3.4.** The output-oriented plant capacity utilization ($PCU_o$) is defined as:

$$PCU_o(x, x^f, y|VRS) = \frac{DF_o(x, y|VRS)}{DF_o^f(x^f, y|VRS)}, \tag{12}$$

where $DF_o(x, y|VRS)$ and $DF_o^f(x^f, y|VRS)$ are output efficiency measures relative to $VRS$ technologies including respectively excluding the variable inputs as defined before. Notice that $0 < PCU_o(x, x^f, y|VRS) \leq 1$, since $1 \leq DF_o(x, y|VRS) \leq DF_o^f(x^f, y|VRS)$. Thus, output-oriented plant capacity utilization has an upper limit of unity, but no lower limit. Following the terminology introduced by Färe, Grosskopf, and Kokkelenberg (1989), Färe, Grosskopf, and Valdmanis (1989) and Färe, Grosskopf, and Lovell (1994) one can distinguish between a so-called biased plant capacity measure $DF_o^f(x^f, y|VRS)$ and an unbiased plant capacity measure $PCU_o(x, x^f, y|VRS)$, where the ratio of efficiency measures ensures to eliminate any existing inefficiency.

Cesaroni, Kerstens, and Van de Woestyne (2017) define a new input-oriented plant capacity measure as follows:

**Definition 3.5.** The input-oriented plant capacity utilization ($PCU_i$) is defined as:

$$PCU_i(x, x^f, y|VRS) = \frac{DF_i^{SR}(x^f, x^v, y|VRS)}{DF_i^{SR}(x^f, x^v, 0|VRS)}, \tag{13}$$

where $DF_i^{SR}(x^f, x^v, y|VRS)$ and $DF_i^{SR}(x^f, x^v, 0|VRS)$ are both sub-vector input efficiency measures reducing only the variable inputs relative to the technology, whereby the latter efficiency measure is evaluated at a zero output level. Notice that $PCU_i(x, x^f, y|VRS) \geq 1$, since $0 < DF_i^{SR}(x^f, x^v, 0|VRS) \leq DF_i^{SR}(x^f, x^v, y|VRS)$. Thus, input-oriented plant capacity utilization has a lower limit of unity, but no upper limit. Similar to the previous case, one can distinguish between a so-called biased plant capacity measure $DF_i^{SR}(x^f, x^v, 0|VRS)$ and an unbiased plant capacity measure $PCU_i^{SR}(x, x^f, y|VRS)$, the latter being cleaned of any prevailing inefficiency.

While these definitions in itself are sufficiently clear, it may be useful to underscore that both these concepts differ with respect to the property of attainability. As stressed by Johansen (1968, p. 362) the output-oriented plant capacity notion is not attainable in that the extra variable inputs necessary to reach the maximal plant capacity output may not be available at the firm level. And even if these extra variable inputs are available at the firm level, restrictions on the available extra variable inputs at the sector level may prevent that all firms simultaneously can reach their maximal plant capacity output. By contrast, the input-oriented plant capacity notion is always attainable in that one can always reduce the amount of existing variable inputs such that one reaches an input set with zero output level. Doing so is possible at the firm level as well as at the sectoral level.

## 3.3 Economic Capacity Concepts: Normalization and Impact of Convexity

Since the literature has abundantly shown that inefficiencies are part and parcel of economic life, following the plant capacity concepts it may be useful to normalize the economic capacity concepts as well. We are inspired by the notion of overall efficiency (see Färe, Grosskopf, and Lovell (1994) or Hackman (2008)), whereby in the case of the cost function one divides the minimal cost by the observed costs ($wx$). Starting from the Definitions 3.1, 3.2 and 3.3, we can now define the normalized economic capacity concepts as follows:

**Definition 3.6.** (i): The normalized minimum of the short-run total cost function $NC(y, w^v, x^f|VRS)$ is $C(y, w^v, x^f|CRS)/wx$.

(ii): Normalized tangency cost with modified fixed inputs $NC^{tang1}(y, w, x^{f*}|VRS)$ is $C(y, w|V)/wx = C(y, w^v, x^{f*}|VRS)/wx$.

(iii): Normalized tangency cost with modified outputs $NC^{tang2}(y(p, w^f, x^f), w, x^f|VRS)$ is $C(y(p, w^f, x^f), w|VRS)/wx = C(y(p, w^f, x^f), w^v, x^f|VRS)/wx$.

(iv): The normalized minimum of long-run total cost function is defined as $NC(y, w^v, x^f|VRS)$ is $C(y, w^v, x^f|CRS)/wx$.

Notice that all of these normalized economic capacity concepts are bounded above at unity, except for the normalized tangency cost with modified outputs $NC^{tang2}(y(p, w^f, x^f), w, x^f|VRS)$ which can be smaller or larger than unity. To understand this phenomenon we must first realize that for observed outputs, we have: $C(y, w|VRS) \neq C(y, w^v, x^f|VRS)$. As a consequence, in Definition 3.2 the optimal tangency cost may be smaller or larger to each of the sides of this inequality. To be explicit, on the one hand we obtain $C(y(p, w^f, x^f), w|VRS) = C(y(p, w^f, x^f), w^v, x^f|VRS) \substack{\geq \\ =\\ <} C(y, w|VRS)$, and on the other hand we get: $C(y(p, w^f, x^f), w|VRS) = C(y(p, w^f, x^f), w^v, x^f|VRS) \substack{\geq \\ =\\ <} C(y, w^v, x^f|VRS)$.

Finally, when comparing convex and nonconvex results, there are cases where plant and economic capacity concepts can be ordered a priori. First, we state these results for the biased plant capacity concepts as well as the non-normalized economic capacity concepts.

**Proposition 3.1.** *(i): For the output-oriented plant capacity utilization, we have:* $DF_o^f(x^f, y|VRS, C) \geq DF_o^f(x^f, y|VRS, NC)$.

*(ii): For the input-oriented plant capacity utilization, we have:* $DF_i^{SR}(x^f, x^v, 0|VRS, C) \leq DF_i^{SR}(x^f, x^v, 0|VRS, NC)$.

*(iii): For the minimum of the short-run total cost function, we have:* $C(y, w^v, x^f|VRS, C) \leq C(y, w^v, x^f|VRS, NC)$.

*(iv): For the tangency cost with modified fixed inputs, we have:* $C^{tang1}(y, w, x^{f*}|VRS, C) \leq C^{tang1}(y, w, x^{f*}|VRS, NC)$.

*(v): For the tangency cost with modified outputs, we have:* $C^{tang2}(y(p, w^f, x^f), w, x^f|VRS, C) \substack{\geq \\ =\\ <} C^{tang2}(y(p, w^f, x^f), w, x^f|VRS, NC)$ .

*(vi): For the minimum of long-run total cost function, we have:* $C(y, w^v, x^f|VRS, C) \leq C(y, w^v, x^f|VRS, NC)$.

Thereafter, we do the same for unbiased plant capacity concepts and the normalized economic capacity concepts.

**Proposition 3.2.** *(i): For the output-oriented plant capacity utilization, we have:* $PCU_o(x, x^f, y|VRS, C) \underset{<}{\overset{\geq}{=}} PCU_o(x, x^f, y|VRS, NC)$.

*(ii): For the input-oriented plant capacity utilization, we have:* $PCU_i(x, x^f, y|VRS, C) \underset{<}{\overset{\geq}{=}} PCU_i(x, x^f, y|VRS, NC)$.

*(iii): For the minimum of the short-run total cost function, we have:* $NC(y, w^v, x^f|VRS, C) \leq NC(y, w^v, x^f|VRS, NC)$.

*(iv): For the tangency cost with modified fixed inputs, we have:* $NC^{tang1}(y, w, x^{f*}|VRS, C) \leq NC^{tang1}(y, w, x^{f*}|VRS, NC)$.

*(v): For the tangency cost with modified outputs, we have:* $NC^{tang2}(y(p, w^f, x^f), w, x^f|VRS, C) \underset{<}{\overset{\geq}{=}} NC^{tang2}(y(p, w^f, x^f), w, x^f|VRS, NC)$ .

*(vi): For the minimum of long-run total cost function, we have:* $NC(y, w^v, x^f|VRS, C) \leq NC(y, w^v, x^f|VRS, NC)$.

Now we are in a position to start developing our empirical illustration.

# 4 Empirical Illustration

## 4.1 Data

To illustrate how the economic and plant capacity notions can be used, we draw upon a secondary data set that is an unbalanced panel of three years (1984-1986) of French fruit producers based on annual accounting data collected in a survey (see Ivaldi, Ladoux, Ossard, and Simioni (1996) for details). Two main criteria determined the selection of 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. 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. The first input capital is considered as fixed.

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). Observe that the limited length of the panel (just three years) justifies the use of an intertemporal frontier accumulating all observations in the technology: this approach fundamentally ignores technical change.

Table 1: Descriptive statistics for French fruit producers (1984-1986)

| Variable | Trimmed mean[a] | Minimum | Maximum |
|---|---|---|---|
| Capital (fixed input) | 85602.58 | 8891 | 500452 |
| Labor (variable input | 229569 | 79569 | 1682201 |
| Materials (variable input) | 157610.9 | 19566 | 1523776 |
| Volume of apple production (output) | 2.146273 | 0.00061 | 37.98153 |
| Volume of other products (output) | 1.37793 | 0.000672 | 25.895 |
| Price of capital | 1.167934 | 0.167802 | 7.889478 |
| Price of labor | 1.059968 | 0.492821 | 1.771435 |
| Price of materials | 6.72676 | 1.732421 | 22.61063 |

Note: [a]10% trimming level.

Table 1 presents basic descriptive statistics for the inputs, the outputs, and the input prices. One observes basically a lot of heterogeneity and a rather wide range for all inputs and outputs. The range for some of the input prices is smaller. More details on the data are available in Ivaldi, Ladoux, Ossard, and Simioni (1996).

In the following, we first discuss the biased plant capacity utilization and non-normalized economic capacity utilization (CU) notions. Thereafter, we study the unbiased plant capacity utilization and normalized economic capacity utilization notions.

## 4.2 Comparing the Biased and Non-Normalized Capacity Utilization Notions

Table 2 shows basic descriptive statistics for all biased and non-normalized capacity utilization notions. We report the average, the standard deviation, and the minima and maxima depending on the context. The relations between convex and nonconvex results are conditioned by the relations described in Proposition 3.1. First, ignoring the CU notion that cannot be ranked (i.e., $NNC^{tang2}$), on average convex and nonconvex results are rather markedly different, except for $BPCU_i$ where the difference is quite small. Second, the range of the results are sometimes different, but some share one of the extremes, except for $BPCU_i$ and $NNC^{tang2}$ for which the range is identical.

Table 2: Descriptive statistics for all biased and non-normalized CU-notions

| **Convex** | $BPCU_o$ | $BPCU_i$ | $NNSRC$ | $NNC^{tang1}$ | $NNC^{tang2}$ | $NNLRC$ |
|---|---|---|---|---|---|---|
| Average | 5.414862 | 0.42333 | 620247.8 | 718839.9 | 315274.8 | 511506.1 |
| Standard deviation | 4.678063 | 0.194978 | 827159.9 | 1124454 | 1058872 | 758764.8 |
| Minimum | 1 | 0.047301 | 10454.19 | 150112.7 | 132380.2 | 8507.063 |
| Maximum | 35.29532 | 1 | 6238552 | 11815722 | 21170527 | 6095270 |
| **Nonconvex** | $BPCU_o$ | $BPCU_i$ | $NNSRC$ | $NNC^{tang1}$ | $NNC^{tang2}$ | $NNLRC$ |
| Average | 2.891018 | 0.430783 | 816915.6 | 1160906 | 301561.7 | 683063.1 |
| Standard deviation | 2.935252 | 0.202152 | 981389.5 | 1730077 | 1043655 | 880893.2 |
| Minimum | 1 | 0.047301 | 14486.9 | 150112.7 | 132380.2 | 13147.43 |
| Maximum | 32.45654 | 1 | 7100639 | 13448388 | 21170527 | 6754195 |

$BPCU_o$: Biased short-run output-oriented plant capacity utilization ($DF_o^f(x^f, y|VRS, .)$).
$BPCU_i$: Biased short-run input-oriented plant capacity utilization ($DF_i^{SR}(x^f, x^v, 0|VRS, .)$).
$NNSRC$: Non-normalized short-run total cost($C(y, w^v, x^f|CRS, .)$).
$NNC^{tang1}$: Non-normalized tangency cost with modified fixed inputs ($C^{tang1}(y, w, x^{f*}|VRS, .)$ ).
$NNC^{tang2}$: Non-normalized tangency cost with modified outputs ($C^{tang2}(y(p, w^f, x^f), w, x^f|VRS, .)$ ).
$NNLRC$: Non-normalized long-run total cost($C(y, w^v, x^f|CRS, .)$ ).

Table 3 reports the results of a formal test statistic proposed by Li (1996) and refined by Fan and Ullah (1999) and Li, Maasoumi, and Racine (2009) lately. The null hypothesis of this Li-test states that both distributions are equal for a given efficiency score or cost frontier estimate and for a given underlying specification of technology. The alternative hypothesis is simply that both distributions are different. This test is valid for both dependent and independent variables. Note that dependency is a characteristic of frontier estimators: frontier efficiency and cost levels depend on sample size, among others.

Table 3: Li test between the biased PCU and non-normalized cost frontier concepts.

| Variables | $BPCU_o$ | $BPCU_i$ | $NNSRC$ | $NNC^{tang1}$ | $NNC^{tang2}$ | $NNLRC$ |
|---|---|---|---|---|---|---|
| $BPCU_o$ | **24.798*** ** | *115.112**** | *288.454**** | *289.826* | *296.06**** | *288.455**** |
| $BPCU_i$ | 173.884*** | **-1.437*** | *288.455**** | *176.102**** | *296.06**** | *176.052**** |
| $NNSRC$ | 288.609*** | 174.2108*** | **3.701*** ** | *16.523**** | *79.0476**** | *0.489* |
| $NNC^{tang1}$ | 133.985*** | 174.24*** | 32.1426*** | **10.798*** ** | *104.128**** | *27.855**** |
| $NNC^{tang2}$ | 295.933*** | 174.268*** | 66.876*** | 67.016*** | **-2.485*** ** | *80.181**** |
| $NNLRC$ | 288.541*** | 288.543*** | 3.873*** | 59.186*** | 76.095*** | **5.925*** ** |

Li test: critical values at 1% level= 2.33($***$); 5% level= 1.64($**$); 10%level= 1.28($*$).

Table 3 is structured as follows. First, components on the diagonal (in bold) depict the Li-test statistic between the convex and nonconvex cases. Second, the components under the diagonal show the Li-test statistic between convex CUs, and the components above the diagonal show the Li-test statistic between nonconvex CUs. The following three conclusions emerge from studying Table 3. First, for the convex capacity notions (below the diagonal) all capacity concepts follow two by two significantly different distributions. Second, for the nonconvex capacity notions (above the diagonal) almost all capacity concepts follow two by two significantly different distributions, except $NNSRC$ and $NNLRC$ that have indistinguishable distributions. Third, all capacity notions follow different distributions under convexity compared to nonconvexity, though the Li-test statistic is only marginally significant for $BPCU_i$ at the 10 % level.

Table 4 reports the Spearman rank correlation coefficients for biased and non-normalized capacity utilization notions. This table is structured in a similar way as Table 3. In this table, components on the diagonal (in bold) depict the rank correlation between the convex and nonconvex cases. The components under the diagonal show the rank correlation between convex CUs and the components above the diagonal show the rank correlation between nonconvex CUs.

Table 4: Spearman rank correlations between the biased PCU and non-normalized cost frontier concepts.

| Variables | $BPCU_o$ | $BPCU_i$ | $NNSRC$ | $NNC^{tang1}$ | $NNC^{tang2}$ | $NNLRC$ |
|---|---|---|---|---|---|---|
| $BPCU_o$ | **0.918**** | *0.304*** | *-0.631*** | *-0.543*** | *-0.106** | *-0.643*** |
| $BPCU_i$ | 0.271** | **0.996**** | *-0.694*** | *-0.707*** | *-0.404*** | *-0.701*** |
| $NNSRC$ | -0.498** | -0.708** | **0.967**** | *0.939*** | *0.469*** | *0.975*** |
| $NNC^{tang1}$ | -0.543** | -0.697** | 0.934** | **0.965**** | *0.579*** | *0.947*** |
| $NNC^{tang2}$ | -0.122* | -0.379** | 0.479** | 0.641** | **0.981**** | *0.462*** |
| $NNLRC$ | -0.646** | -0.684** | 0.960** | 0.950** | 0.460** | **0.988**** |

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

The following three conclusions emerge from studying Table 4. First, for the convex results, one can observe that $BPCU_i$ rank correlates better with all cost-based CU notions in absolute values than $BPCU_o$, and that $NNSRC$ and $NNLRC$ have the highest rank correlation among cost-based CU notions. Second, for the nonconvex results, exactly the same two conclusions emerge. Third, comparing convex and nonconvex results, the rank correlations are remarkably high overall, and these are highest for $BPCU_i$ and lowest for $BPCU_o$.

## 4.3 Comparing the Unbiased and Normalized Capacity Utilization Notions

Table 5 shows basic descriptive statistics for all unbiased and normalized capacity utilization notions. We report the average, the standard deviation, and the minima and maxima depending on the context. In this case, the relations between convex and nonconvex results are determined by the relations described in Proposition 3.2. First, ignoring the three CU notions that cannot be ranked, on average convex and nonconvex results are rather markedly different for the three other CU notions (i.e., $NSRC$, $NC^{tang1}$ and $NLRC$). Second, the range of the results differ sometimes. But, some share one of the extremes, except for $PCU_i$ and $NC^{tang2}$ for which the range is again identical.

Table 5: Descriptive statistics for all CU-notions

| **Convex** | $PCU_o$ | $PCU_i$ | $NSRC$ | $NC^{tang1}$ | $NC^{tang2}$ | $NLRC$ |
|---|---|---|---|---|---|---|
| Average | 0.710459 | 1.733724 | 0.331119 | 0.43418 | 0.391276 | 0.260619 |
| Standard deviation | 0.221112 | 1.636011 | 0.173495 | 0.189827 | 2.280111 | 0.161353 |
| Minimum | 0.070056 | 1 | 0.053327 | 0.103932 | 0.017937 | 0.036357 |
| Maximum | 1 | 21.14141 | 1 | 1 | 45.5131 | 1 |
| **Nonconvex** | $PCU_o$ | $PCU_i$ | $NSRC$ | $NC^{tang1}$ | $NC^{tang2}$ | $NLRC$ |
| Average | 0.690958 | 2.539953 | 0.464158 | 0.629439 | 0.372487 | 0.378417 |
| Standard deviation | 0.244674 | 2.156585 | 0.253099 | 0.247589 | 2.270101 | 0.218591 |
| Minimum | 0.096771 | 1 | 0.069012 | 0.133735 | 0.017937 | 0.039328 |
| Maximum | 1 | 21.14141 | 1 | 1 | 45.5131 | 1 |

$PCU_o^{SR}$: Unbiased short-run output-oriented plant capacity utilization ($PCU_i^{SR}(x, x^f, y|VRS, .)$).
$PCU_i^{SR}$: Unbiased short-run input-oriented plant capacity utilization ($PCU_o^{SR}(x, x^f, y|VRS, .)$).
$NSRC$: Normalized short-run total cost ($C(y, w^v, x^f|CRS, .)/wx$).
$NC^{tang1}$: Normalized tangency cost with modified fixed inputs ($C^{tang1}(y, w, x^{f*}|VRS, .)/wx$).
$NC^{tang2}$: Normalized tangency cost with modified outputs ($C^{tang2}(y(p, w^f, x^f), w, x^f|VRS, .)/wx$).
$NLRC$: Normalized long-run total cost($C(y, w^v, x^f|CRS, .)/wx$).

Table 6 reports the Li-test statistics and it is structured in a similar way as Table 3 above. A glance at Table 6 yields the following conclusions. First, for the convex capacity notions (below the diagonal) almost all capacity concepts follow two by two significantly different distributions, except $NC^{tang2}$ and $NLRC$ that have indistinguishable distributions. Second, for the nonconvex capacity notions (above the diagonal) all capacity concepts follow two by two significantly different distributions. Third, all capacity notions follow different distributions under convexity compared to nonconvexity.

Table 6: Li test for all unbiased CUs

| Variables | $PCU_o$ | $PCU_i$ | $NSRC$ | $NC^{tang1}$ | $NC^{tang2}$ | $NLRC$ |
|---|---|---|---|---|---|---|
| $PCU_o$ | **10.251*** ** | *96.205**** | *25.171**** | *4.083**** | *83.440**** | *49.933**** |
| $PCU_i$ | 74.606*** | **31.005*** ** | *105.896**** | *93.404**** | *143.246**** | *116.441**** |
| $NSRC$ | 48.981*** | 160.506*** | **6.621*** ** | *9.953**** | *20.426**** | *3.725**** |
| $NC^{tang1}$ | 30.828*** | 157.123*** | 13.382*** | **25.034*** ** | *58.883**** | *27.888**** |
| $NC^{tang2}$ | 78.834*** | 148.022*** | 17.429*** | 51.662*** | **-2.852*** ** | *14.522**** |
| $NLRC$ | 65.281*** | 163.884*** | 10.122*** | 43.685*** | 0.851 | **12.632*** ** |

Li test: critical values at 1% level= 2.33($***$); 5% level= 1.64($**$); 10%level= 1.28($*$).

Table 7 reports the Spearman rank correlation coefficients for unbiased and normalized capacity utilization notions. In this table, the components on the diagonal show the rank correlation between convex and nonconvex case. The components under the diagonal show the rank correlation between convex CUs and the components above the diagonal show the rank correlation between nonconvex CUs.

Table 7: Spearman rank correlations for all unbiased CUs

| Variables | $PCU_o$ | $PCU_i$ | $NSRC$ | $NC^{tang1}$ | $NC^{tang2}$ | $NLRC$ |
|---|---|---|---|---|---|---|
| $PCU_o$ | **0.706** ** | *0.418*** | *0.253*** | *0.042* | *-0.403*** | *0.224*** |
| $PCU_i$ | 0.181** | **0.888** ** | *0.443*** | *0.307*** | *-0.661*** | *0.521*** |
| $NSRC$ | -0.241** | 0.650** | **0.893** ** | *0.734*** | *0.041* | *0.935*** |
| $NC^{tang1}$ | -0.295** | 0.156** | 0.569** | **0.807** ** | *0.273*** | *0.769*** |
| $NC^{tang2}$ | -0.326** | -0.534** | -0.025 | 0.630** | **0.997** ** | *0.016* |
| $NLRC$ | -0.084 | 0.749** | 0.894** | 0.585** | -0.074 | **0.957** ** |

** Correlation is significant at the 0.01 level(2-tailed).

* Correlation is significant at the 0.05 level(2-tailed).

For the convex results, one can notice that $PCU_i$ rank correlates better with all cost-based CU notions in absolute values than $PCU_o$, except for the $NC^{tang1}$ CU notion. Furthermore, $NSRC$ and $NLRC$ again obtain the highest rank correlation among cost-based CU notions. Finally, $PCU_o$ essentially has a zero correlation with $NLRC$. For the nonconvex results, exactly the same two conclusions emerge (but now without any exception). In addition, $PCU_o$ has now a close to zero correlation with $NC^{tang1}$. Comparing convex and nonconvex results, the rank correlations are still high overall, and these are highest now for $NC^{tang2}$ and again lowest for $PCU_o$.

# 5 Conclusions

This contribution has set itself two major goals. A first major goal has been to make a theoretically coherent input-oriented comparison between these technical and economic capacity notions. As a point of comparison, also the output-oriented plant capacity notion has been included. A second major goal has been to make this coherent input-oriented comparison among capacity notions using both convex and nonconvex technologies to assess the impact of the convexity axiom. Theoretically, the investigation of this convexity hypothesis has led us to establish the cases where plant and economic capacity concepts can be ordered a priori (see Propositions 3.1 and 3.2).

The empirical results have shown the following key results. First, there appears quite some heterogeneity among the different technical and economic capacity notions in terms of descriptive statistics. Second, formal testing has revealed that in almost all cases technical and economic capacity notions follow different distributions. Thus, each of these concepts seems to capture a different part of economic reality. Furthermore, each and every capacity concept seems also to follow almost always a different distribution under convexity and nonconvexity. Thus, convexity matters from a distributional viewpoint. Third, the study of Spearman rank correlation coefficients reveals that almost uniformly the input-oriented plant capacity notion correlates better with the cost-based capacity notions than the output-oriented plant capacity notion. Furthermore, the rank correlations are overall high for convex and nonconvex results. Thus, convexity seems to matter less from a ranking point of view.

Therefore, two key conclusions emerge from this contribution. First, the recently introduced input-oriented plant capacity notion lends itself more naturally to comparisons with cost-based capacity notions than the more traditional output-oriented plant capacity notion. Thus, while the output-oriented plant capacity notions enjoys some popularity in empirical applications (see the literature review in Cesaroni, Kerstens, and Van de Woestyne (2017)), applied researchers should probably consider using the new input-oriented plant capacity notion that is more in line with the traditional cost-based capacity notions widespread in economics.

Second, convexity matters also for both technical and economic capacity notions. Therefore, it seems essential to further empirically explore potential differences between estimates based on convex and nonconvex technologies and cost functions in even greater detail (e.g., the impact on economies of scope, the effect on mergers and acquisitions, etc.). Thus, even though theoretically the impact of convexity is known since a while, it is important to further explore the impact of convexity on key economic value relations in practice. The current evidence provided shows that this impact is nonnegligible when measuring capacity and that convexification may not be harmless.

As an agenda for future research, we can mention three issues. First, it would be good if our empirical results regarding both the comparison of input-oriented technical and economic capacity notions as well as the impact of the convexity axiom in this context would be corroborated in additional empirical work by other researchers. Second, while the input-oriented plant capacity notion compares well with cost-based capacity notions, one may wonder whether the traditional output-oriented plant capacity would fit much better with capacity notions based on the revenue function (see, e.g., Lindebo, Hoff, and Vestergaard (2007) or Segerson and Squires (1995)). This conjecture remains to be explored. Third, the fact that the output-oriented plant capacity notion is not attainable while the new input-oriented plant capacity notion satisfies attainability merits further investigation. It is an open question whether and how the output-oriented plant capacity notion can be made attainable.

# References


Anderson, R., D. Lewis, and T. Springer (2000): "Operating Efficiencies in Real Estate: A Critical Review of the Literature," *Journal of Real Estate Literature*, 8(1), 1–18.

Briec, W., K. Kerstens, D. Prior, and I. Van de Woestyne (2010): "Tangency Capacity Notions Based upon the Profit and Cost Functions: A Non-Parametric Approach and a Comparison," *Economic Modelling*, 27(5), 1156–1166.

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.

Cassels, J. M. (1937): "Excess Capacity and Monopolistic Competition," *Quarterly Journal of Economics*, 51(3), 426–443.

Cesaroni, G., K. Kerstens, and I. Van de Woestyne (2017): "A New Input-Oriented Plant Capacity Notion: Definition and Empirical Comparison," *Pacific Economic Review*, 22(4), 720–739.

Christiano, L. (1981): "A Survey of Measures of Capacity Utilization," *IMF Staff Papers*, 28(1), 144–198.

Coelli, T., E. Grifell-Tatjé, and S. Perelman (2002): “Capacity Utilisation and Profitability: A Decomposition of Short-Run Profit Efficiency,” *International Journal of Production Economics*, 79(3), 261–278.

Cummins, D., and M. Weiss (2000): “Analyzing Firm Performance in the Insurance Industry Using Frontier Efficiency and Productivity Methods,” in *Handbook of Insurance*, ed. by G. Dionne, pp. 767–829. Kluwer, Boston.

De Borger, B., K. Kerstens, D. Prior, and I. Van de Woestyne (2012): “Static Efficiency Decompositions and Capacity Utilisation: Integrating Technical and Economic Capacity Notions,” *Applied Economics*, 44(31), 4125–4141.

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.

Färe, R., S. Grosskopf, and E. Kokkelenberg (1989): “Measuring Plant Capacity, Utilization and Technical Change: A Nonparametric Approach,” *International Economic Review*, 30(3), 655–666.

Färe, R., S. Grosskopf, and C. Lovell (1994): *Production Frontiers*. Cambridge University Press, Cambridge.

Färe, R., S. Grosskopf, and V. Valdmanis (1989): “Capacity, Competition and Efficiency in Hospitals: A Nonparametric Approach,” *Journal of Productivity Analysis*, 1(2), 123–138.

Giménez, V., and D. Prior (2007): “Long- and Short-Term Cost Efficiency Frontier Evaluation: Evidence from Spanish Local Governments,” *Fiscal Studies*, 28(1), 121–139.

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

Harker, P., and S. Zenios (2001): “What Drives the Performance of Financial Institutions?,” in *Performance of Financial Institutions: Efficiency, Innovation, Regulation*, ed. by P. Harker, and S. Zenios, pp. 3–31. Cambridge University Press, Cambridge.

Hatami-Marbini, A., A. Emrouznejad, and M. Tavana (2011): “A Taxonomy and Review of the Fuzzy Data Envelopment Analysis Literature: Two Decades in the Making,” *European Journal of Operational Research*, 214(3), 457–472.

Hickman, B. G. (1964): “On a New Method of Capacity Estimation,” *Journal of the American Statistical Association*, 59(306), 529–549.

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.

Jacobsen, S. (1970): "Production Correspondences," *Econometrica*, 38(5), 754–771.

Jamasb, T., and M. Pollitt (2000): "Benchmarking and Regulation: International Electricity Experience," *Utilities Policy*, 9(3), 107–130.

Johansen, L. (1968): "Production functions and the concept of capacity," *Recherches récentes sur la fonction de production, Collection, Economie mathématique et économétrie*, 2, 52.

Kerstens, K., and I. Van de Woestyne (2014): "Solution Methods for Nonconvex Free Disposal Hull Models: A Review and Some Critical Comments," *Asia-Pacific Journal of Operational Research*, 31(1).

Klein, L. (1960): "Some Theoretical Issues in the Measurement of Capacity," *Econometrica*, 28(2), 272–286.

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

Li, Q., E. Maasoumi, and J. Racine (2009): "A Nonparametric Test for Equality of Distributions with Mixed Categorical and Continuous Data," *Journal of Econometrics*, 148(2), 186–200.

Lindebo, E. (2005): "Multi-national Industry Capacity in the North Sea Flatfish Fishery," *Marine Resource Economics*, 20(4), 385–406.

Lindebo, E., A. Hoff, and N. Vestergaard (2007): "Revenue-based Capacity Utilisation Measures and Decomposition: The Case of Danish North Sea Trawlers," *European Journal of Operational Research*, 180(1), 215–227.

Luh, Y.-H., and S. Stefanou (1991): "Productivity Growth in U.S. Agriculture under Dynamic Adjustment.," *American Journal of Agricultural Economics*, 73(4), 1116–1125.

Nelson, R. A. (1989): "On the Measurement of Capacity Utilization," *Journal of Industrial Economics*, 37(3), 273–286.

Ozcan, Y. (2008): *Health Care Benchmarking and Performance Evaluation: An Assessment using Data Envelopment Analysis (DEA)*. Springer, Berlin.

Prior-Jiménez, D. (2003): "Long- and Short-Run Non-Parametric Cost Frontier Efficiency: An Application to Spanish Savings Banks," *Journal of Banking & Finance*, 27(4), 655–671.

Sahoo, B., and K. Tone (2009): "Decomposing Capacity Utilization in Data Envelopment Analysis: An Application to Banks in India," *European Journal of Operational Research*, 195(2), 575–594.

Segerson, K., and D. Squires (1990): "On the Measurement of Economic Capacity Utilization for Multi-Product Industries," *Journal of Econometrics*, 44(3), 347–361.

Segerson, K., and D. Squires (1995): "Measurement of Capacity Utilization for Revenue-Maximizing Firms," *Bulletin of Economic Research*, 47(1), 77–85.

Sherman, H., and G. Ladino (1995): "Managing Bank Productivity Using Data Envelopment Analysis (DEA)," *Interfaces*, 25(2), 60–73.

Thanassoulis, E., M. Silva Portela, and O. Despić (2008): "DEA - The Mathematical Programming Approach to Efficiency Analysis," in *The Measurement of Productive Efficiency and Productivity Change*, ed. by H. Fried, C. Lovell, and S. Schmidt, pp. 251–420. Oxford University Press, New York.

Tone, K., and B. Sahoo (2003): "Scale, Indivisibilities and Production Function in Data Envelopment Analysis," *International Journal of Production Economics*, 84(2), 165–192.

Voigt, S. (2016): "Determinants of Judicial Efficiency: A Survey," *European Journal of Law and Economics*, 42(2), 183–208.

Walden, J., J. Kirkley, and A. Kitts (2003): "A Limited Economic Assessment of the Northeast Groundfish Fishery Buyout Program," *Land Economics*, 79(3), 426–439.

Walden, J., and D. Tomberlin (2010): "Estimating Fishing Vessel Capacity: A Comparison of Nonparametric Frontier Approaches," *Marine Resource Economics*, 25(1), 23–36.

Worthington, A. (2001): "An Empirical Survey of Frontier Efficiency Measurement Techniques in Education," *Education Economics*, 9(3), 245–268.

# Appendix: Computing Plant and Economic Capacity Notions (Supplementary Material)

This appendix is devoted to show how in a non-parametric frontier framework the components of of the various capacity concepts can be estimated. To specify the models for the estimation, we first recall the notation introduced in this contribution. The vector of $N$ inputs ($x \in \mathbb{R}_+^N$) allows producing a vector of $M$ outputs ($y \in \mathbb{R}_+^M$). The vector of input prices is given by $w \in \mathbb{R}_+^N$. These vectors of inputs and input prices can be partitioned into a fixed and variable part (denoted $x = (x^v, x^f)$ and $w = (w^v, w^f)$). Assume that for each of the $K$ observed production units that need to be evaluated ($k = 1, \ldots, K$), we know both the vector of $M$ outputs obtained from the vector of $N$ inputs as well as the corresponding input prices. Let $Y$ denote the $(K \times M)$ matrix of observed outputs and let $X$ denote the $(K \times N)$ matrix of observed inputs. Elements of these matrices can be denoted as follows: for the $k$-th producer, the quantity of the inputs is denoted $x_k$ with a corresponding input price $w_k$, while the quantity of the outputs is denoted by $y_k$. The corresponding fixed and variable input components are denoted by $x_k^f$ and $x_k^v$. Finally, since non-parametric frontier technologies are founded on activity analysis, we need a vector of activity variables $z = (z_1, \ldots, z_K)$ indicating the intensity levels at which each of these $K$ observed activities is conducted.

## Computing Plant and Economic Capacity Notions: Convex Case

The estimation of the plant and economic capacity components under convexity by using non-parametric frontier methods implies the resolution of the following series of linear programming problems for each observation $(x_o, y_o)$.

The output-oriented radial technical efficiency measure $DF_o(x, y|VRS)$ is computed by optimizing the following linear program:

$$\begin{array}{ll}
\max & \theta \\
s.t & \sum\limits_{k=1}^{K} z_k y_k \geq \theta y_o, \\
& \sum\limits_{k=1}^{K} z_k x_k \leq x_o, \\
& \sum\limits_{k=1}^{K} z_k = 1, \\
& \theta \geq 0, z_k \geq 0, \quad k = 1, \ldots, K.
\end{array} \tag{14}$$

The short-run output-oriented radial technical efficiency measure $DF_o^f(x^f, y|VRS)$ requires the optimization of the following linear program:

$$\begin{aligned} \max \quad & \theta \\ s.t \quad & \sum_{k=1}^{K} z_k y_k \geq \theta y_o, \\ & \sum_{k=1}^{K} z_k x_k^f \leq x_o^f, \\ & \sum_{k=1}^{K} z_k = 1, \\ & \theta \geq 0, z_k \geq 0, \quad k = 1, \ldots, K. \end{aligned} \tag{15}$$

The input efficiency measure reducing only the variable inputs $(DF_i^{SR}(x^f, x^v, y|VRS))$ is computed by optimizing the following linear program:

$$\begin{aligned} \min \quad & \lambda \\ s.t \quad & \sum_{k=1}^{K} z_k y_k \geq y_o, \\ & \sum_{k=1}^{K} z_k x_k^f \leq x_o^f, \\ & \sum_{k=1}^{K} z_k x_k^v \leq \lambda x_o^v, \\ & \sum_{k=1}^{K} z_k = 1, \\ & \theta \geq 0, z_k \geq 0, \quad k = 1, \ldots, K. \end{aligned} \tag{16}$$

The input-oriented short-run efficiency measure reducing variable inputs evaluated relative to the input set with a zero output level $(DF_i^{SR}(x^f, x^v, 0|VRS))$ is computed by optimizing the following linear program:

$$\begin{aligned} \min \quad & \lambda \\ s.t \quad & \sum_{k=1}^{K} z_k y_k \geq 0, \\ & \sum_{k=1}^{K} z_k x_k^f \leq x_o^f, \\ & \sum_{k=1}^{K} z_k x_k^v \leq \lambda x_o^v, \\ & \sum_{k=1}^{K} z_k = 1, \\ & \theta \geq 0, z_k \geq 0, \quad k = 1, \ldots, K. \end{aligned} \tag{17}$$

The minimum of the short-run total cost function $C(y, w^v, x^f|VRS)$ is computed as $w_o^f x^f + C(y, w^v, x^f|CRS)$ where:

$$
\begin{aligned}
C(y, w^v, x^f|CRS) = \quad \min \quad & w_o^v x^v \\
s.t \quad & \sum_{k=1}^{K} z_k y_k \geq y_o, \\
& \sum_{k=1}^{K} z_k x_k^f \leq x_o^f, \\
& \sum_{k=1}^{K} z_k x_k^v \leq x^v, \\
& x^v \geq 0, z_k \geq 0, \quad k = 1, \dots, K.
\end{aligned} \tag{18}
$$

To obtain the tangency cost with modified fixed inputs $C^{tang1}(y, w, x^{f*}|VR)$, one needs to find the optimal value for the following linear programming problem:

$$
\begin{aligned}
C(y, w|VRS) = \quad \min \quad & w_o x \\
s.t \quad & \sum_{k=1}^{K} z_k y_k \geq y_o, \\
& \sum_{k=1}^{K} z_k x_k \leq x, \\
& \sum_{k=1}^{K} z_k = 1, \\
& x \geq 0, z_k \geq 0, \quad k = 1, \dots, K.
\end{aligned} \tag{19}
$$

To obtain the tangency cost with modified outputs $C^{tang2}(y(p, w^f, x^f), w, x^f|VRS)$, one can solve the following model:

$$
\begin{aligned}
C(y(p, w^f, x^f), w^v, w^f|VRS) = \quad \min \quad & w_o^v x^v + w_{opt}^f x_o^f \\
s.t \quad & \sum_{k=1}^{K} z_k y_k \geq y_{opt}, \\
& \sum_{k=1}^{K} z_k x_k \leq x^v, \\
& \sum_{k=1}^{K} z_k x_k \leq x_o^f, \\
& \sum_{k=1}^{K} z_k = 1, \\
& x^v \geq 0, z_k \geq 0. \quad k = 1, \dots, K.
\end{aligned} \tag{20}
$$

where $(y_{opt}, w_{opt}^f)$ is obtained by solving the following nonlinear system of inequalities (see

Briec, Kerstens, Prior, and Van de Woestyne (2010) for details).

$$\begin{cases} p.y_k - w_o^v.x_k^v - w_{opt}^f.x_k^f - (p.y_{opt} - w_o^v.x^v - w_{opt}^f.x_o^f) \leq 0, \quad k = 1, ..., K, \\ \sum_{k=1}^{K} z_k y_k \geq y_{opt}, \\ \sum_{k=1}^{K} z_k x_k \leq x^v, \\ \sum_{k=1}^{K} z_k x_k \leq x_o^f, \\ \sum_{k=1}^{K} z_k = 1, \\ y_{opt} \geq 0, x^v \geq 0, p \geq 0, w_{opt}^f \geq 0, z_k \geq 0, \qquad k = 1, \ldots, K. \end{cases} \tag{21}$$

To obtain the minimum of long run total cost function $C(y, w|VRS)$, we require finding the optimal value of the following linear programming problem:

$$\begin{aligned} C(y, w|CRS) = \quad \min \quad & w_o x \\ s.t \quad & \sum_{k=1}^{K} z_k y_k \geq y_o, \\ & \sum_{k=1}^{K} z_k x_k \leq x, \\ & z_k \geq 0, x \geq 0. \end{aligned} \tag{22}$$

## Computing Plant and Economic Capacity Notions: Nonconvex Case

The estimation of the plant and economic capacity components under nonconvexity by using non-parametric frontier methods implies the resolution of the following series of mathematical programming problems for each observation $(x_o, y_o)$.

The output-oriented radial technical efficiency measure $DF_o(x, y|VRS)$ is computed by optimizing the following binary mixed integer program:

$$\begin{aligned} \max \quad & \theta \\ s.t \quad & \sum_{k=1}^{K} z_k y_k \geq \theta y_o, \\ & \sum_{k=1}^{K} z_k x_k \leq x_o, \\ & \sum_{k=1}^{K} z_k = 1, \\ & \theta \geq 0, z_k \in \{0, 1\}, \quad k = 1, \ldots, K. \end{aligned} \tag{23}$$

The output-oriented short-run radial technical efficiency measure $DF_o^f(x^f, y|VRS)$ requires the optimization of the following model:

$$
\begin{aligned}
\max \quad & \theta \\
s.t \quad & \sum_{k=1}^{K} z_k y_k \geq \theta y_o, \\
& \sum_{k=1}^{K} z_k x_k^f \leq x_o^f, \\
& \sum_{k=1}^{K} z_k = 1, \\
& \theta \geq 0, z_k \in \{0,1\}, \quad k = 1, \ldots, K.
\end{aligned} \tag{24}
$$

The input efficiency measure reducing only the variable inputs $(DF_i^{SR}(x^f, x^v, y|VRS))$ is computed by optimizing the following binary mixed integer program:

$$
\begin{aligned}
\min \quad & \lambda \\
s.t \quad & \sum_{k=1}^{K} z_k y_k \geq y_o, \\
& \sum_{k=1}^{K} z_k x_k^f \leq x_o^f, \\
& \sum_{k=1}^{K} z_k x_k^v \leq \lambda x_o^v, \\
& \sum_{k=1}^{K} z_k = 1, \\
& \theta \geq 0, z_k \in \{0,1\}, \quad k = 1, \ldots, K.
\end{aligned} \tag{25}
$$

The input-oriented short run efficiency measure reducing variable inputs evaluated relative to the input set with a zero output level $(DF_i^{SR}(x^f, x^v, 0|VRS))$ is computed by optimizing the following binary mixed integer program:

$$
\begin{aligned}
\min \quad & \lambda \\
s.t \quad & \sum_{k=1}^{K} z_k y_k \geq 0, \\
& \sum_{k=1}^{K} z_k x_k^f \leq x_o^f, \\
& \sum_{k=1}^{K} z_k x_k^v \leq \lambda x_o^v, \\
& \sum_{k=1}^{K} z_k = 1, \\
& \theta \geq 0, z_k \in \{0,1\}, \quad k = 1, \ldots, K.
\end{aligned} \tag{26}
$$

The minimum of short-run total cost function $C(y, w^v, x^f|VRS)$ is computed as $w_o^f x^f + C(y, w^v, x^f|CRS)$ in which:

$$
\begin{aligned}
C(y, w^v, x^f|CRS) = \quad \min \quad & w_o^v x^v \\
s.t \quad & \sum_{k=1}^{K} \alpha z_k y_k \geq y_o, \\
& \sum_{k=1}^{K} \alpha z_k x_k^f \leq x_o^f, \\
& \sum_{k=1}^{K} \alpha z_k x_k^v \leq x^v, \\
& \sum_{k=1}^{K} z_k = 1, \\
& x^v \geq 0, \alpha \geq 0, z_k \in \{0,1\}, \quad k = 1, \ldots, K.
\end{aligned}
\tag{27}
$$

To obtain the tangency cost with modified fixed inputs $C^{tang1}(y, w, x^{f*}|VRS)$, one needs to find the optimal value of the following binary mixed integer programming problem:

$$
\begin{aligned}
C(y, w|VRS) = \quad \min \quad & w_o x \\
s.t \quad & \sum_{k=1}^{K} z_k y_k \geq y_o, \\
& \sum_{k=1}^{K} z_k x_k \leq x, \\
& \sum_{k=1}^{K} z_k = 1, \\
& x \geq 0, z_k \in \{0,1\}, \quad k = 1, \ldots, K.
\end{aligned}
\tag{28}
$$

To obtain the tangency cost with modified outputs $C^{tang2}(y(p, w^f, x^f), w, x^f|VRS)$, one can solve the following model.

$$
\begin{aligned}
C(y(p, w^f, x^f), w^v, w^f|VRS) = \quad \min \quad & w_o^v x^v + w_{opt}^f x_o^f \\
s.t \quad & \sum_{k=1}^{K} z_k y_k \geq y_{opt}, \\
& \sum_{k=1}^{K} z_k x_k \leq x^v, \\
& \sum_{k=1}^{K} z_k x_k \leq x_o^f, \\
& \sum_{k=1}^{K} z_k = 1, \\
& x^v \geq 0, z_k \in \{0,1\}, \quad k = 1, \ldots, K,
\end{aligned}
\tag{29}
$$

where $(y_{opt}, w^f_{opt})$ is obtained by solving the following nonlinear binary mixed integer system of inequalities (an extension of Briec, Kerstens, Prior, and Van de Woestyne (2010)).

$$\begin{cases} p.y_k - w^v_o.x^v_k - w^f_{opt}.x^f_k - (p.y_{opt} - w^v_o.x^v - w^f_{opt}.x^f_o) \leq 0, \quad k = 1, ..., K, \\ \sum_{k=1}^{K} z_k y_k \geq y_{opt}, \\ \sum_{k=1}^{K} z_k x_k \leq x^v, \\ \sum_{k=1}^{K} z_k x_k \leq x^f_o, \\ \sum_{k=1}^{K} z_k = 1, \\ y_{opt} \geq 0, x^v \geq 0, p \geq 0, w^f_{opt} \geq 0, z_k \in \{0,1\} \qquad k = 1, \ldots, K. \end{cases} \tag{30}$$

To obtain the minimum of the long-run total cost function $C(y, w|VRS)$, we require finding the optimal value of the following binary mixed integer programming problem:

$$\begin{aligned} C(y, w|CRS) = \quad & \min \quad w_o x \\ & s.t \quad \sum_{k=1}^{K} \alpha z_k y_k \geq y_o, \\ & \qquad \sum_{k=1}^{K} \alpha z_k x_k \leq x, \\ & \qquad \sum_{k=1}^{K} z_k = 1, \\ & \qquad \alpha \geq 0, x \geq 0, z_k \in \{0,1\}, \ k = 1, ..., K. \end{aligned} \tag{31}$$