Estimation and Inference in Nonparametric Frontier Models: Recent Developments and Perspectives

Estimation and Inference in Nonparametric Frontier Models: Recent Developments and Perspectives

Léopold Simar

Université Catholique de Louvain
Belgium
leopold.simar@uclouvain.be

Paul W. Wilson

Clemson University
USA
pww@clemson.edu

now
the essence of knowledge
Boston – Delft

Foundations and Trends® in Econometrics

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is L. Simar and P. W. Wilson, Estimation and Inference in Nonparametric Frontier Models: Recent Developments and Perspectives, Foundations and Trends® in Econometrics, vol 5, nos 3–4, pp 183–337, 2011

ISBN: 978-1-60198-666-5

Foundations and Trends® in Econometrics

Volume 5 Issues 3–4, 2011

Editorial Scope

Foundations and Trends® in Econometrics will publish survey and tutorial articles in the following topics:

- Econometric Models:
  - Identification
  - Model Choice and Specification Analysis
  - Non-linear Regression Models
- Simultaneous Equation Models
- Estimation Frameworks
- Biased Estimation
- Computational Problems
- Microeconometrics
- Treatment Modeling
- Discrete Choice Modeling
- Models for Count Data
- Duration Models
- Limited Dependent Variables
- Panel Data
- Time Series Analysis:
  - Dynamic Specification
  - Inference and Causality
  - Continuous Time Stochastic Models
  - Modeling Non-linear Time Series
  - Unit Roots
  - Cointegration
- Latent Variable Models
- Qualitative Response Models
- Hypothesis Testing
- Econometric Theory:
  - Interactions-based Models
  - Duration Models
- Financial Econometrics
- Measurement Error in Survey Data
- Productivity Measurement and Analysis
- Semiparametric and Nonparametric Estimation
- Bootstrap Methods
- Nonstationary Time Series
- Robust Estimation

Information for Librarians

Foundations and Trends® in Econometrics, 2011, Volume 5, 4 issues. ISSN paper version 1551-3076. ISSN online version 1551-3084. Also available as a combined paper and online subscription.

Foundations and Trends® in
Econometrics
Vol. 5, Nos. 3–4 (2011) 183–337

DOI: 10.1561/0800000020

Estimation and Inference in Nonparametric Frontier Models: Recent Developments and Perspectives

Léopold Simar[1] and Paul W. Wilson[2]

[1] *Institut de Statistique, Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Voie du Roman Pays 20, B 1348 Louvain-la-Neuve, Belgium, leopold.simar@uclouvain.be*
[2] *Department of Economics and School of Computing, 228 Sirrine Hall, Clemson University, Clemson, South Carolina, 29634–1309, USA, pww@clemson.edu*

Abstract

Nonparametric estimators are widely used to estimate the productive efficiency of firms and other organizations, but often without any attempt to make statistical inference. Recent work has provided statistical properties of these estimators as well as methods for making statistical inference, and a link between frontier estimation and extreme value theory has been established. New estimators that avoid many of the problems inherent with traditional efficiency estimators have also been developed; these new estimators are robust with respect to outliers and avoid the well-known curse of dimensionality. Statistical properties, including asymptotic distributions, of the new estimators

have been uncovered. Finally, several approaches exist for introducing environmental variables into production models; both two-stage approaches, in which estimated efficiencies are regressed on environmental variables, and conditional efficiency measures, as well as the underlying assumptions required for either approach, are examined.

Contents

1

Nonparametric Statistical Models of Production: Combining Economics and Statistics

The economic theory underlying analysis of efficiency in production dates at least to the work of Koopmans (1951), Debreu (1951), and Farrell (1957). Farrell made the first attempt to estimate efficiency from a set of observed production units, but the statistical properties of his estimator were only considered much later.

The discussion that follows introduces basic concepts and notation; *Economic Considerations* introduces an economic framework, to which *Statistical Considerations* adds a statistical paradigm.

1.1 Economic Considerations

Producers transform inputs into outputs; for example, in manufacturing, inputs typically include labor, capital, energy, materials, and perhaps other things, while outputs are the products produced. There may be one, several, perhaps many different products that are produced. Of course, production is constrained by what is possible or feasible. Let $\boldsymbol{x} \in \mathbb{R}_+^p$ and $\boldsymbol{y} \in \mathbb{R}_+^q$ denote vectors of input and output quantities, respectively, and let

$$\mathcal{P} = \{(\boldsymbol{x},\boldsymbol{y}) \mid \boldsymbol{x} \text{ can produce } \boldsymbol{y}\} \tag{1.1}$$

denote the set of feasible combinations of inputs and outputs, i.e., the production set. Any output quantities $\boldsymbol{y}$ can be produced using input quantities $\boldsymbol{x}$ if and only if $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}$. However, the points in $\mathcal{P}$ are not equally desirable.

The following three assumptions regarding $\mathcal{P}$ are standard in microeconomic theory of the firm; see, for example, Shephard (1970) and Färe (1988).

Assumption 1.1. $\mathcal{P}$ is closed.

Assumption 1.2. All production requires use of some inputs: $(\boldsymbol{x},\boldsymbol{y}) \notin \mathcal{P}$ if $\boldsymbol{x} = 0$ and $\boldsymbol{y} \geq \boldsymbol{0}$, $\boldsymbol{y} \neq \boldsymbol{0}$.[1]

Assumption 1.3. Both inputs and outputs are freely disposable: if $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}$, then for any $(\boldsymbol{x}',\boldsymbol{y}')$ such that $\boldsymbol{x}' \geq \boldsymbol{x}$ and $\boldsymbol{y}' \leq \boldsymbol{y}$, $(\boldsymbol{x}',\boldsymbol{y}') \in \mathcal{P}$.

Assumption 1.1 ensures that the boundary of $\mathcal{P}$ is included in $\mathcal{P}$. Assumption 1.2 means that there are no "free lunches." The free disposability assumption is sometimes called strong disposability and is equivalent to an assumption of monotonicity of the technology. This property also characterizes the technical possibility of wasting resources (i.e., the possibility of producing less with more resources).

For purposes of efficiency measurement, the upper boundary of $\mathcal{P}$ is relevant. The efficient subset of points in $\mathcal{P}$ is the upper boundary (frontier) of $\mathcal{P}$, i.e., the locus of optimal production plans (e.g., minimal achievable input level for a given output, or maximal achievable output given the level of the inputs). The upper boundary of $\mathcal{P}$,

$$\mathcal{P}^{\partial} = \{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P} \mid (\gamma^{-1}\boldsymbol{x},\gamma\boldsymbol{y}) \notin \mathcal{P} \; \forall\, \gamma \in (1,\infty)\} \tag{1.2}$$

is sometimes referred to as the *technology* or the *production frontier,* and is given by the intersection of $\mathcal{P}$ and the closure of its complement.

[1] Throughout, inequalities involving vectors are assumed to hold element by element; e.g., $\boldsymbol{a} \leq \boldsymbol{b}$ denotes $a_j \leq b_j$ for each $j = 1, \ldots, k$, where k is the length of $\boldsymbol{a}$ and $\boldsymbol{b}$.

Firms that are technically inefficient operate at points in the interior of $\mathcal{P}$, while those that are technically efficient operate somewhere along the technology defined by $\mathcal{P}^{\partial}$.

Various features of the production set $\mathcal{P}$ and its frontier $\mathcal{P}^{\partial}$ are often of interest to applied researchers. One such feature is *returns to scale*. Strictly speaking, returns to scale is a feature of the frontier, $\mathcal{P}^{\partial}$, but it is common to ascribe such features to the set $\mathcal{P}$. There are several possibilities.

Definition 1.1. The frontier $\mathcal{P}^{\partial}$ displays globally constant returns to scale (CRS) if and only if $(\alpha\boldsymbol{x},\alpha\boldsymbol{y}) \in \mathcal{P}\ \forall\ (\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}$ and $\alpha \in [0,\infty)$.

Definition 1.2. The frontier $\mathcal{P}^{\partial}$ displays globally variable returns to scale (VRS) if and only if for any $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}$, there exist constants $a(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}^1_+$, $b(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}^1_{++}$ such that $a(\boldsymbol{x},\boldsymbol{y}) \leq b(\boldsymbol{x},\boldsymbol{y})$ and $(\alpha\boldsymbol{x},\alpha\boldsymbol{y}) \in \mathcal{P}\ \forall\ \alpha \in [a(\boldsymbol{x},\boldsymbol{y}),\ b(\boldsymbol{x},\boldsymbol{y})]$.

Definition 1.3. The frontier $\mathcal{P}^{\partial}$ displays globally nonincreasing returns to scale (NIRS) if and only if $(\alpha\boldsymbol{x},\alpha\boldsymbol{y}) \in \mathcal{P}\ \forall\ (\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}$ and $\alpha \in [0,1]$.

Note that Definition 1.2 encompasses Definitions 1.1 and 1.3; i.e., CRS and NIRS are special cases of VRS. In the same way, Definition 1.3 encompasses Definition 1.1 in that CRS is a special case of NIRS. In other words, assuming either CRS or NIRS is more restrictive than assuming VRS; assuming CRS is more restrictive that assuming NIRS.

The production set $\mathcal{P}$ can also be described by its sections or level sets. For instance, the input requirement set for some $\boldsymbol{y} \in \mathbb{R}^q_+$ is given by

$$\mathcal{X}(\boldsymbol{y}) = \{\boldsymbol{x} \in \mathbb{R}^p_+ \mid (\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}\}, \tag{1.3}$$

i.e., the set of all input vectors $\boldsymbol{x}$ that can produce output vector $\boldsymbol{y}$. The boundary of this set, i.e., the (input-oriented) efficiency boundary $\mathcal{X}^{\partial}(y)$, is defined for a given $y \in \mathbb{R}^q_+$ by

$$\mathcal{X}^{\partial}(\boldsymbol{y}) = \{\boldsymbol{x} \mid \boldsymbol{x} \in \mathcal{X}(\boldsymbol{y}), \theta\boldsymbol{x} \notin \mathcal{X}(\boldsymbol{y}),\ \forall\ \theta \in (0,1)\}. \tag{1.4}$$

Alternatively, the output feasibility set for some $\boldsymbol{x} \in \mathbb{R}^p_+$ is defined by

$$\mathcal{Y}(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{R}^q_+ \mid (\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}\}, \tag{1.5}$$

which gives the set of all output vectors $\boldsymbol{y}$ than can be produced with given input quantities $\boldsymbol{x}$. The (output-oriented) efficiency boundary $\mathcal{Y}^\partial(\boldsymbol{x})$ is defined, for a given $\boldsymbol{x} \in \mathbb{R}^p_+$, as

$$\mathcal{Y}^\partial(\boldsymbol{x}) = \{\boldsymbol{y} \mid \boldsymbol{y} \in \mathcal{Y}(\boldsymbol{x}), \lambda\boldsymbol{y} \notin \mathcal{Y}(\boldsymbol{x}),\ \forall\, \lambda > 1\}. \tag{1.6}$$

Then the production set $\mathcal{P}$ corresponds to the union of all sets $\mathcal{X}(\boldsymbol{y})$ over all $\boldsymbol{y} \in \mathbb{R}^q_+$, or to the union of all sets $\mathcal{Y}(\boldsymbol{x})$ over all $\boldsymbol{x} \in \mathbb{R}^p_+$.

The Debreu–Farrell input measure of *technical* efficiency for a given point $(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}^{p+q}_+$ is given by

$$\begin{aligned}\theta(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P}) &= \inf\{\theta \mid \theta\boldsymbol{x} \in \mathcal{X}(\boldsymbol{y})\}\\ &= \inf\{\theta \mid (\theta\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}\}.\end{aligned} \tag{1.7}$$

Note that this measure is defined for some points in $\mathbb{R}^{p+q}_+$ not necessarily in $\mathcal{P}$ (i.e., points for which a solution exists in (1.7)). Given an output level y, and an input mix (a direction) given by the vector $\boldsymbol{x}$, the corresponding efficient level of input is given by

$$\boldsymbol{x}^\partial(\boldsymbol{y}) = \theta(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})\boldsymbol{x}, \tag{1.8}$$

which is the projection of $(\boldsymbol{x},\boldsymbol{y})$ onto the efficient boundary $\mathcal{P}^\partial$, along the ray $\boldsymbol{x}$ and orthogonal to the vector $\boldsymbol{y}$.

Figure 1.1 illustrates a point $(\boldsymbol{x}_0,\boldsymbol{y}_0) \in \mathcal{P}$ for $p = q = 1$. The level set $\boldsymbol{x}^\partial(\boldsymbol{y})$ defined in (1.8) contains just one point in Figure 1.1; in terms of the labels on the horizontal axis, $\theta(\boldsymbol{x}_0,\boldsymbol{y}_0 \mid \mathcal{P}) = \boldsymbol{x}^\partial(\boldsymbol{y}_0)/\boldsymbol{x}_0 < 1$.

In general, for $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}$, $\theta(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})$ gives the feasible proportionate reduction of inputs that a unit located at $(\boldsymbol{x},\boldsymbol{y})$ could undertake to become technically efficient. By construction, for all $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}$, $\theta(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P}) \in (0,1]$; $(\boldsymbol{x},\boldsymbol{y})$ is technically efficient if and only if $\theta(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P}) = 1$. This measure is the reciprocal of the Shephard (1970) input distance function.

Similarly, in the output direction, the Debreu–Farrell output measure of technical efficiency is given by

$$\begin{aligned}\lambda(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P}) &= \sup\{\lambda \mid \lambda\boldsymbol{y} \in \mathcal{Y}(\boldsymbol{x})\}\\ &= \sup\{\lambda \mid (\boldsymbol{x},\lambda\boldsymbol{y}) \in \mathcal{P}\}\end{aligned} \tag{1.9}$$

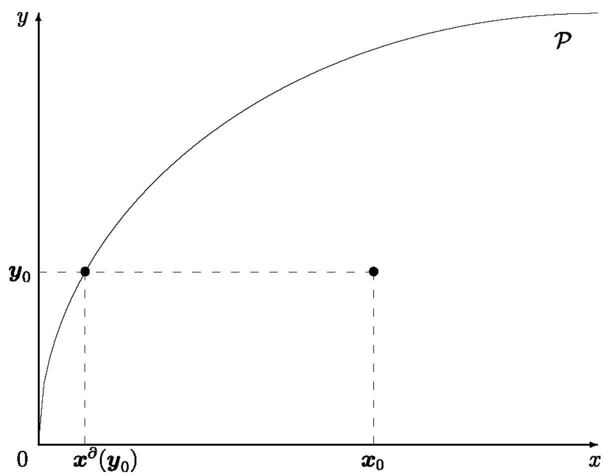

Fig. 1.1 Input efficiency measure.

for $(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}_+^{p+q}$. Analogous to the input-oriented case described above, $\lambda(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})$ gives the feasible proportionate increase in outputs for a unit operating at $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}$ that would achieve technical efficiency. By construction, for all $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}$, $\lambda(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P}) \in [1,\infty)$ and $(\boldsymbol{x},\boldsymbol{y})$ is technically efficient if and only if $\lambda(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P}) = 1$.

The output efficiency measure $\lambda(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})$ is the reciprocal of the Shephard (1970) output distance function. The efficient level of output, for the input level x and for the direction of the output vector determined by y, is given by

$$\boldsymbol{y}^{\partial}(\boldsymbol{x}) = \lambda(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})\boldsymbol{y}. \tag{1.10}$$

Figure 1.2 illustrates the same point $(\boldsymbol{x}_0,\boldsymbol{y}_0) \in \mathcal{P}$ shown in Figure 1.1. Here, the set $\boldsymbol{y}^{\partial}(\boldsymbol{x}_0)$ defined by (1.10) also contains a single point, and in terms of the labels on the vertical axis in Figure 1.2, $\lambda(\boldsymbol{x}_0,\boldsymbol{y}_0 \mid \mathcal{P}) = \boldsymbol{y}^{\partial}(\boldsymbol{x}_0)/\boldsymbol{y}_0 > 1$.

Efficiency can be measured in other directions, although care should be taken to avoid having efficiency measures depend on units of measurement for inputs or outputs. For example, a hyperbolic measure of efficiency is given by

$$\gamma(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P}) = \sup\{\gamma \mid (\gamma^{-1}\boldsymbol{x},\gamma\boldsymbol{y}) \in \mathcal{P}\} \tag{1.11}$$

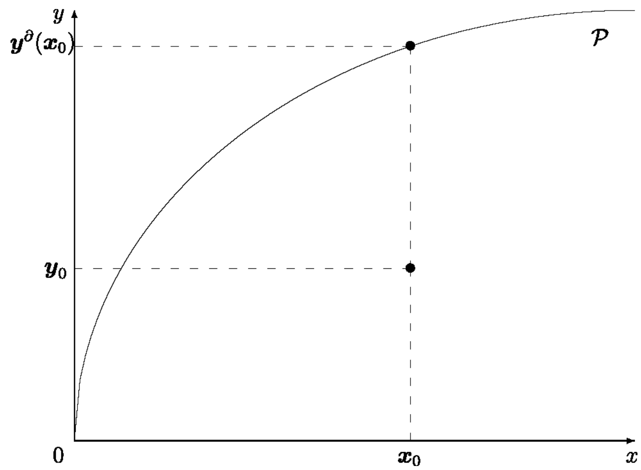

Fig. 1.2 Output efficiency measure.

for $(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}_+^{p+q}$. This hyperbolic measure of efficiency gives the simultaneous proportionate, feasible reduction in input levels and the proportionate, feasible increase in output levels for a unit operating at $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}$ that would result in technical efficiency, and is the reciprocal of the hyperbolic graph measure of efficiency defined by Färe et al. (1985).

In terms of the illustration in Figure 1.3, the firm operating at the point $(\boldsymbol{x}_0,\boldsymbol{y}_0) \in \mathcal{P}$ can become technically efficient by moving along the curved (hyperbolic) path from $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ to $(\boldsymbol{x}_\gamma^{\partial}(\boldsymbol{x}_0,\boldsymbol{y}_0),\ \boldsymbol{y}_\gamma^{\partial}(\boldsymbol{x}_0,\boldsymbol{y}_0))$. By construction, $\gamma(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P}) \in [1,\infty)$ for all $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{P}$; in addition, $(\boldsymbol{x},\boldsymbol{y})$ is technically efficient if and only if $\gamma(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P}) = 1$. For $(\boldsymbol{x},\boldsymbol{y})$ in the interior of $\mathcal{P}$, the corresponding hyperbolic-efficient levels of inputs are outputs that are given by

$$\left(\boldsymbol{x}_\gamma^{\partial}(\boldsymbol{x},\boldsymbol{y}),\ \boldsymbol{y}_\gamma^{\partial}(\boldsymbol{x},\boldsymbol{y})\right) = \left(\gamma(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})^{-1}\boldsymbol{x},\ \gamma(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})\boldsymbol{y}\right). \tag{1.12}$$

Alternatively, Chambers et al. (1996) introduced the directional efficiency measure defined by

$$\delta(\boldsymbol{x},\boldsymbol{y} \mid \boldsymbol{u},\boldsymbol{v},\mathcal{P}) = \sup\{\delta \mid (\boldsymbol{x} - \delta\boldsymbol{u}, \boldsymbol{y} + \delta\boldsymbol{v}) \in \mathcal{P}\}, \tag{1.13}$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ are direction vectors with $\boldsymbol{u} \in \mathbb{R}_+^p$, $\boldsymbol{v} \in \mathbb{R}_+^q$, and $\left[\boldsymbol{u}' \quad \boldsymbol{v}'\right] \neq \boldsymbol{0}$. This distance function projects a point $(\boldsymbol{x},\boldsymbol{y})$ onto the

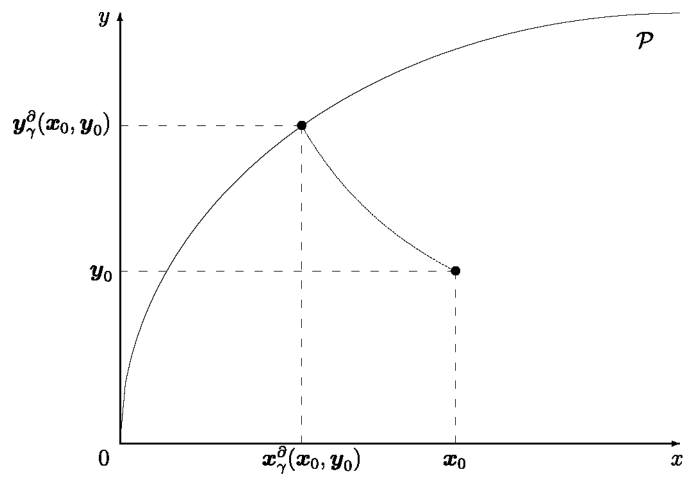

Fig. 1.3 Hyperbolic efficiency measure.

frontier $\mathcal{P}^\partial$ in the direction $(-\boldsymbol{u}, \boldsymbol{v})$, with

$$\left(\boldsymbol{x}_\delta^\partial(\boldsymbol{x}, \boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{v}),\ \boldsymbol{y}_\delta^\partial(\boldsymbol{x}, \boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{v})\right) = \\ (\boldsymbol{x} - \delta(\boldsymbol{x}, \boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{v}, \mathcal{P})\boldsymbol{u},\ \boldsymbol{y} + \delta(\boldsymbol{x}, \boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{v}, \mathcal{P})\boldsymbol{v}) \tag{1.14}$$

giving the directionally-efficient (in the direction $(\boldsymbol{u}, \boldsymbol{v})$) levels of inputs and outputs.

The directional distance function is illustrated in Figure 1.4. Setting $\boldsymbol{u} = \boldsymbol{x}_0$, $\boldsymbol{v} = \boldsymbol{y}_0$, the firm operating at $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ becomes technically efficient when it moves to $(\boldsymbol{x}_\delta^\partial(\boldsymbol{x}_0, \boldsymbol{y}_0),\ \boldsymbol{y}_\delta^\partial(\boldsymbol{x}_0, \boldsymbol{y}_0))$ defined by (1.14) (the notation indicating dependence on direction vectors is suppressed here and in Figure 1.4 to conserve space). By construction, $\delta(\boldsymbol{x}, \boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{v}, \mathcal{P}) \in [0, \infty)$ for all $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{P}$; a point $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{P}$ is technically efficient if and only if $\delta(\boldsymbol{x}, \boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{v}, \mathcal{P}) = 0$.

Färe et al. (2008, p. 534) state that the directional distance function is independent of units of measurement in the sense that

$$\delta(\boldsymbol{\alpha}_x \circ \boldsymbol{x}, \boldsymbol{\alpha}_y \circ \boldsymbol{y} \mid \boldsymbol{\alpha}_x \circ \boldsymbol{u}, \boldsymbol{\alpha}_y \circ \boldsymbol{v}, \mathcal{P}) = \delta(\boldsymbol{x}, \boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{v}, \mathcal{P}), \tag{1.15}$$

where $\boldsymbol{\alpha}_x \in \mathbb{R}_{++}^p$, $\boldsymbol{\alpha}_y \in \mathbb{R}_{++}^q$, and $\circ$ denotes the Hadamard product.[2] However, while (1.15) is true, it also indicates that if units of

[2] The Hadamard product of two arrays $A = [a_{ij}]$ and $B = [b_{ij}]$ with the same dimensions is given by the array $C = [c_{ij}]$ having the same dimensions as A and B where $c_{ij} = a_{ij}b_{ij}$;

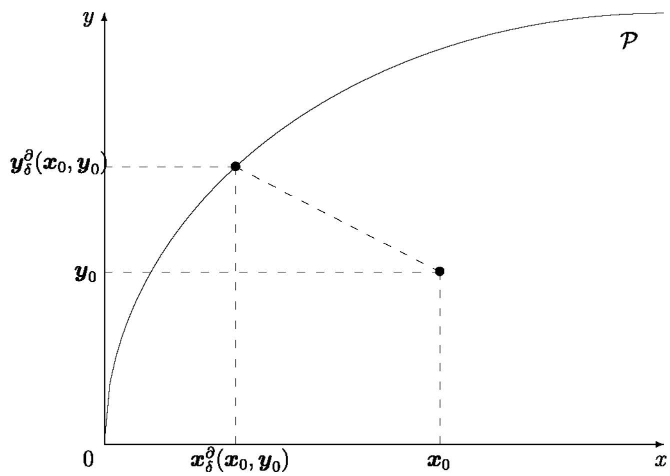

Fig. 1.4 Directional efficiency measure.

measurement for inputs or outputs are changed, the corresponding direction vector must be rescaled to avoid changing the value of the directional distance function. Instead of being homogeneous of degree zero with respect to inputs and outputs, the directional distance function is only homogeneous of degree zero with respect to inputs, outputs, *and direction vectors.*

This feature of the directional distance function makes the range of reasonable choices for the direction vectors less broad than has been suggested in the literature. For example, Färe et al. (2008, p. 533) note that the direction vectors should be specified in the same units as the inputs and outputs, but then go on to suggest choosing $\boldsymbol{u} = 1$, $\boldsymbol{v} = 1$ or to optimize $\boldsymbol{u}$ and $\boldsymbol{v}$ to minimize distance to the (estimated) frontier. But, if one specifies $\boldsymbol{u} = 1$, $\boldsymbol{v} = 1$, and then changes the units of measurement, this will require re-scaling also $\boldsymbol{u}$ and $\boldsymbol{v}$ so that they no longer equal unity in order to avoid changing the value of the distance function. Hence the choice of $(1, 1)$ for $(\boldsymbol{u}, \boldsymbol{v})$ is arbitrary, and therefore rather meaningless. Moreover, if the direction vectors are optimized to

e.g., see Marcus and Kahn (1959), Marcus and Thompson (1963), and Johnson (1974a, 1974b).

minimize distance to the estimated frontier, then the results will be sensitive to the units of measurement that are used.

It is easy to show that if $\boldsymbol{u} = \boldsymbol{x}$ and $\boldsymbol{v} = 0$, then

$$\delta(\boldsymbol{x}, \boldsymbol{y} \mid \boldsymbol{u} = \boldsymbol{x}, \boldsymbol{v} = 0, \mathcal{P}) = 1 - \theta(\boldsymbol{x}, \boldsymbol{y} \mid \mathcal{P}). \tag{1.16}$$

Similarly, if $\boldsymbol{u} = 0$ and $\boldsymbol{v} = \boldsymbol{y}$, then

$$\delta(\boldsymbol{x}, \boldsymbol{y} \mid \boldsymbol{u} = 0, \boldsymbol{v} = \boldsymbol{y}, \mathcal{P}) = \lambda(\boldsymbol{x}, \boldsymbol{y} \mid \mathcal{P}) - 1. \tag{1.17}$$

A common choice, when $\boldsymbol{x} \in \mathbb{R}^p_{++}$ and $\boldsymbol{y} \in \mathbb{R}^q_{++}$, is to set $\boldsymbol{u} = \boldsymbol{x}$ and $\boldsymbol{v} = \boldsymbol{y}$. One can also set the direction vectors equal to the sample means of inputs and outputs in order to use a common direction for all observations. For additional properties of the directional distance function, see Chambers et al. (1996).

Both the hyperbolic and the directional measures are measures of *technical* efficiency, as are the input- and output-oriented measures discussed above. Technical efficiency refers to what is possible, but as suggested earlier, not everything that is possible is desirable. Firms may want to maximize profits, which requires considering the prices of inputs and outputs in addition to their quantities. In the case of government provision of goods and services, or in regulated industries, producers' goals may be cost minimization or perhaps revenue maximization.

A variety of assumptions on $\mathcal{P}$ are found in the literature (e.g., free disposability, convexity, etc.; see Shephard, 1970 for examples). The assumptions about $\mathcal{P}$ determine the appropriate estimator that should be used to estimate $\mathcal{P}^\partial$, $\theta(\boldsymbol{x}, \boldsymbol{y} \mid \mathcal{P})$, $\lambda(\boldsymbol{x}, \boldsymbol{y} \mid \mathcal{P})$, $\gamma(\boldsymbol{x}, \boldsymbol{y})$, or $\delta(\boldsymbol{x}, \boldsymbol{y})$. This issue will be discussed next.

1.2 Statistical Considerations

In real-world research problems, the attainable set $\mathcal{P}$, as well as $\mathcal{X}(\boldsymbol{y})$, $\mathcal{X}^\partial(\boldsymbol{y})$, $\mathcal{Y}(\boldsymbol{x})$, and $\mathcal{Y}^\partial(\boldsymbol{x})$ are unknown to the analyst. Consequently, the efficiency scores $\theta(\boldsymbol{x}, \boldsymbol{y} \mid \mathcal{P})$, $\lambda(\boldsymbol{x}, \boldsymbol{y} \mid \mathcal{P})$, $\gamma(\boldsymbol{x}, \boldsymbol{y} \mid \mathcal{P})$, and $\delta(\boldsymbol{x}, \boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{v}, \mathcal{P})$ corresponding to a particular unit operating at $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{P}$ are also unknown.

In ordinary settings, the only information available to the researcher is a sample

$$\mathcal{S}_n = \{(\boldsymbol{X}_i, \boldsymbol{Y}_i),\ i = 1,\ \ldots,\ n\} \tag{1.18}$$

of observations on input and output levels for a set of production units engaged in the activity of interest.[3] The statistical paradigm raises the following question that must be answered: what can be learned by observing $\mathcal{S}_n$? In other words, how can the information in $\mathcal{S}_n$ be used to estimate $\mathcal{P}$, $\theta(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})$, $\lambda(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})$ $\gamma(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})$, $\delta(\boldsymbol{x},\boldsymbol{y} \mid \boldsymbol{u},\boldsymbol{v},\mathcal{P})$, or other things of interest?

Answering these questions involves much more than reading the data in $\mathcal{S}_n$ into a computer program and pushing some buttons on the keyboard to solve some linear programs. A relevant question is, "what is learned from an estimate of $\theta(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})$, $\lambda(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})$, or other numbers computed from $\mathcal{S}_n$?" The answer is clear and certain: *almost nothing.* One might learn, for example, that unit A uses less input quantities while producing greater output quantities than unit B, but little else can be learned from estimates of the efficiency measures introduced above without doing some additional work.

Before anything can be learned about $\theta(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})$, $\lambda(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})$, $\gamma(\boldsymbol{x},\boldsymbol{y} \mid \mathcal{P})$, $\delta(\boldsymbol{x},\boldsymbol{y} \mid \boldsymbol{u},\boldsymbol{v},\mathcal{P})$, or by extension about $\mathcal{P}$ and its various features, one must use methods of statistical analysis to understand the properties of whatever estimators have been used to obtain estimates of the things of interest.[4] This raises the following questions: Is the estimator *consistent*? Is the estimator *biased*? If the estimator is biased, does the bias disappear as the sample size tends toward infinity? If the estimator is biased, can the bias be corrected, and at what cost; i.e., does correcting the bias introduce too much noise? Can confidence intervals for the values of interest be estimated, and if so, how? How might one test interesting hypotheses about the production process? Notions of statistical consistency, etc. are discussed below.

[3] Following standard notation, random variables are denoted by upper-case letters, and realizations of random variables and other nonstochastic quantities by lower-case letters.

[4] Note that an *estimator* is a random variable, while an *estimate* is a realization of an estimator (random variable). An estimator can take perhaps infinitely many values with different probabilities, while an estimate is merely a known, nonrandom value.

Before these questions can be answered, a statistical model must be defined; without a statistical model, one cannot know what is estimated. Statistical models consist of two parts: (i) a probability model, which in the present case includes assumptions on the production set $\mathcal{P}$ and the distribution of input and output vectors $(\boldsymbol{x}, \boldsymbol{y})$ over $\mathcal{P}$; and (ii) a sampling model describing how data are obtained from the probability model. The statistical model provides a theoretical description the mechanism that yields the data in the sample $\mathcal{S}_n$, and is sometimes called the data-generating process (DGP). In typical research settings, the task is to use the data in $\mathcal{S}_n$ to learn something about the features of the DGP.

In cases where a group of productive units are observed at the same point in time, i.e., where cross-sectional data are observed, it is convenient and often reasonable to assume the sampling process involves independent draws from the probability distribution defined in the DGP's probability model. With regard to the probability model, one must attempt reasonable assumptions. Of course, there are trade-offs here; the assumptions on the probability model must be strong enough to permit estimation using estimators that have useful properties, and to allow those properties to be deduced, yet not so strong as to impose conditions on the DGP that do not reflect reality. The goal should be, in all cases, to make minimal, flexible assumptions in order to let the data reveal as much as possible about the underlying DGP, as opposed to making strong, untested assumptions that might influence the results of estimation and inference in perhaps large and misleading ways. The assumptions defining the statistical model are of crucial importance, since any inference that might be made will typically be valid only if the assumptions are in fact *true.*

The above considerations apply equally to parametric as well as nonparametric approaches to estimation and inference. One can imagine a spectrum of estimation approaches, ranging from fully parametric (most restrictive) to fully nonparametric (least restrictive). Fully parametric estimation strategies necessarily involve stronger assumptions on the probability model, which is completely specified in terms of a specific probability distribution function, structural equations, etc. Semi-parametric strategies are less restrictive; in these approaches, some (but not all) features of the probability model are left unspecified (for example, in a regression setting one might specify parametric forms for some, but not all, of the moments of a distribution function in the probability model). Fully nonparametric approaches assume no parametric forms for any features of the probability model. Instead, only (relatively) mild assumptions on broad features of the probability distribution are made, usually involving assumptions of various types of continuity, degrees of smoothness, etc.

With fully nonparametric approaches to efficiency estimation, no specific analytical function describing the frontier is assumed. In addition, possibly restrictive assumptions on the stochastic part of the model, describing the probabilistic behavior of the observations in the sample with respect to the efficient boundary of $\mathcal{P}$, are also avoided. There is, however, a cost for this flexibility; in particular, all observed input–output pairs $(\boldsymbol{X}_i, \boldsymbol{Y}_i)$ are assumed to be technically attainable; observations $(\boldsymbol{X}_i, \boldsymbol{Y}_i)$ on input, output vectors are assumed to be drawn randomly and independently from a *population* of firms whose input–output vectors are distributed on the attainable set $\mathcal{P}$ according to some unknown probability law described by a probability density function $f(\boldsymbol{x}, \boldsymbol{y})$ or the corresponding distribution function $F(\boldsymbol{x}, \boldsymbol{y}) = \Pr(\boldsymbol{X} \leq \boldsymbol{x}, \boldsymbol{Y} \leq \boldsymbol{y})$, with

$$\Pr(\boldsymbol{X}_i, \boldsymbol{Y}_i) \in \mathcal{P}) = 1. \tag{1.19}$$

By contrast, fully parametric approaches to efficiency estimation developed by Aigner et al. (1977), Meeusen and van den Broeck (1977), Battese and Corra (1977), Jondrow et al. (1982), and others allow some observations to lie outside the production set $\mathcal{P}$ by incorporating a (two-sided) stochastic term reflecting measurement error or other noise in addition to a (one-sided) stochastic term reflecting inefficiency. Introduction of the stochastic noise term, however, incurs a cost: some parametric structure is required for such models to be identified, which in turn requires assumptions that may or may not be supported by data. In addition, such models typically allow for only a single response variable, i.e., a single output variable in a production framework; researchers typically work in a cost framework when there are multiple outputs, but this in turn requires data on input-prices.

To the extent that the fully parametric approach allows for measurement error, it only does so for the response variable, and not for any of the explanatory variables. Perhaps most problematic, writing the model in a regression framework introduces issues of causality and exogeneity that do not arise in the fully nonparametric approach, which more closely resembles an exercise in density estimation.

The fully parametric approaches are often called *stochastic frontier analysis*, while the fully nonparametric approaches are frequently called *deterministic frontier analysis.* This terminology is unfortunate since it is misleading and has created a good bit of confusion in the literature. In both approaches, there is only one frontier, and it is fixed, not stochastic. In both approaches, the location of the frontier is unknown, and this is what necessitates estimation and gives rise to uncertainty. In both approaches, the distance from a given observation $(\boldsymbol{X}_i, \boldsymbol{Y}_i)$ to the frontier (in any direction) is unknown, and must be estimated.

The most popular nonparametric efficiency estimators are based on the idea of estimating the attainable set $\mathcal{P}$ by the smallest set $\widehat{\mathcal{P}}$ within some class of sets that envelop the observed data. Depending on assumptions made on $\mathcal{P}$, this idea leads to the Free Disposal Hull (FDH) estimator of Deprins et al. (1984), which relies only on an assumption of free disposability, and the Data Envelopment Analysis (DEA) estimators which incorporate additional assumptions. Farrell (1957) was the first to use a DEA estimator in an empirical application, but the idea remained obscure until it was popularized by Charnes et al. (1978) and Banker et al. (1984). Charnes et al. estimated $\mathcal{P}$ by the convex cone of the FDH estimator of $\mathcal{P}$, thus imposing an assumption of constant returns to scale, while Banker et al. used the convex hull of the FDH estimator of $\mathcal{P}$, thereby allowing for VRS.

The primary advantage of nonparametric models and estimators lies in their great flexibility (as opposed to parametric, deterministic frontier models). In addition, the nonparametric estimators are easy to compute, and today most of their statistical properties are well-established. As will be discussed below, inference is available using bootstrap methods.

The main drawbacks of the fully nonparametric DEA and FDH estimators is that they are very sensitive to outliers and extreme values, and that noisy data are not allowed. Fortunately, robust alternatives to DEA and FDH estimators are available for use in fully nonparametric models; these alternative approaches will be described later. Also, as discussed below in the last section of this survey, "stochastic" versions of DEA and FDH estimators are the object of current research.

It should be noted that allowing for noise in frontier models presents difficult problems even in a fully parametric framework where one can rely on the assumed parametric structure. In fully parametric models where the DGP involves a one-sided error process reflecting inefficiency and a two-sided error process reflecting statistical noise, numerical identification of the statistical model's features is sometimes highly problematic even with large (but finite) samples; see Ritter and Simar (1997) for examples.

Apart from the issue of numerical identification, fully parametric frontier models that incorporate a noise term present other difficulties. Efficiency estimates in these models are based on residual terms that are unidentified. Researchers instead base efficiency estimates on an expectation, conditional on a composite residual; estimating an *expected inefficiency* is rather different from estimating *actual* inefficiency. An additional problem arises from the fact that, even if the fully parametric, stochastic frontier model is correctly specified, there is typically a nontrivial probability of drawing samples with the "wrong" skewness (e.g., when estimating cost functions, one would expect composite residuals with right-skewness, but it is certainly possible to draw finite samples with left-skewness — the probability of doing so depends on the sample size and the mean of the composite errors). Since there are apparently no published studies, and also apparently no working papers in circulation, where researchers report composite residuals with the "wrong" skewness when fully parametric, stochastic frontier models are estimated, it appears that estimates are sometimes, perhaps often, conditioned (i) on either drawing observations until the desired skewness is obtained or (ii) on model specifications that result in the desired skewness. This raises formidable questions for inference; see Simar and Wilson (2010) for discussion.

References

Aigner, D., C. A. K. Lovell, and P. Schmidt (1977), 'Formulation and estimation of stochastic frontier production function models'. *Journal of Econometrics* **6**, 21–37.

Aragon, Y., A. Daouia, and C. Thomas-Agnan (2005), 'Nonparametric frontier estimation: A conditional quantile-based approach'. *Econometric Theory* **21**, 358–389.

Athreya, B. K. (1987), 'Bootstrap of the mean in the infinite variance case'. *Annals of Statistics* **15**, 724–731.

Banker, D. R. (1993), 'Maximum likelihood, consistency and data envelopment analysis: A statistical foundation'. *Management Science* **39**, 1265–1273.

Banker, D. R., A. Charnes, and W. W. Cooper (1984), 'Some models for estimating technical and scale inefficiencies in data envelopment analysis'. *Management Science* **30**, 1078–1092.

Banker, D. R. and R. Natarajan (2008), 'Evaluating contextual variables affecting productivity using data envelopment analysis'. *Operations Research* **56**, 48–58.

Battese, E. G. and G. S. Corra (1977), 'Estimation of a production frontier model: With application to the pastoral zone off eastern Australia'. *Australian Journal of Agricultural Economics* **21**, 169–179.

Beran, R. and G. Ducharme (1991), *Asymptotic Theory for Bootstrap Methods in Statistics*. Montreal: Centre de Reserches Mathematiques, University of Montreal.

Beran, R. and M. S. Srivastava (1985), 'Bootstrap tests and confidence regions for functions of a covariance matrix'. *Annals of Statistics* **13**, 95–115.

Bickel, J. P. and D. A. Freedman (1981), 'Some asymptotic theory for the bootstrap'. *Annals of Statistics* **9**, 1196–1217.

Bickel, J. P. and A. Sakov (2008), 'On the choice of m in the m out of n bootstrap and confidence bounds for extrema'. *Statistica Sinica* **18**, 967–985.

Bogetoft, P. (1996), 'DEA on relaxed convexity assumptions'. *Managment Science* **42**, 457–465.

Bogetoft, P., J. M. Tama, and J. Tind (2000), 'Convex input and output projections of nonconvex production possibility sets'. *Management Science* **46**, 858–869.

Bretagnolle, J. (1983), 'Lois limites du bootstrap des certaines fonctionnelles'. *Annales de l'Institut Henri Poincare (Section B)* **19**, 223–234.

Briec, W., K. Kerstens, and P. Van den Eeckaut (2004), 'Non-convex technologies and cost functions: Definitions, duality, and nonparametric tests of convextiy'. *Journal of Economics* **81**, 155–192.

Bădin, L., C. Daraio, and L. Simar (2010), 'Optimal bandwidth selection for conditional efficiency measures: A data-driven approach'. *European Journal of Operational Research* **201**, 633–664.

Bădin, L., C. Daraio, and L. Simar (2012a), 'Explaining inefficiency in nonparametric production models: The state of the art'. *Annals of Operations Research*, Forthcoming.

Bădin, L., C. Daraio, and L. Simar (2012b), 'How to measure the impact of environmental factors in a nonparametric production model'. *European Journal of Operational Research* **223**, 818–833.

Bădin, L. and L. Simar (2009), 'A bias-corrected nonparametric envelopment estimator of frontiers'. *Econometric Theory* **25**, 1289–1318.

Cameron, C. A. and P. K. Trivedi (2005), *Microeconometrics: Methods and Applications*. Cambridge: Cambridge University Press.

Cazals, C., J. P. Florens, and L. Simar (2002), 'Nonparametric frontier estimation: A robust approach'. *Journal of Econometrics* **106**, 1–25.

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

Charnes, A., W. W. Cooper, and E. Rhodes (1978), 'Measuring the efficiency of decision making units'. *European Journal of Operational Research* **2**, 429–444.

Charnes, A., W. W. Cooper, and E. Rhodes (1981), 'Evaluating program and managerial efficiency: An application of data envelopment analysis to program follow through'. *Management Science* **27**, 668–697.

Coelli, T., D. S. P. Rao, and G. E. Battese (1997), *An Introduction to Efficiency and Productivity Analysis*. Boston: Kluwer Academic Publishers.

Cook, D. R. and S. Weisberg (1982), *Residuals and Influence in Regression*. New York: Chapman and Hall.

Daouia, A., J. P. Florens, and L. Simar (2008), 'Functional convergence of quantile-type frontiers with application to parametric approximations'. *Journal of Statistical Planning and Inference* **138**, 708–725.

Daouia, A., J. P. Florens, and L. Simar (2010), 'Frontier estimation and extreme value theory'. *Bernoulli* **16**, 1039–1063.

Daouia, A., J. P. Florens, and L. Simar (2012), 'Regularization of non-parametric frontier estimators'. *Journal of Econometrics* **168**, 285–299.

Daouia, A. and I. Gijbels (2011a), 'Estimating frontier cost models using extremiles'. In: I. Van Keilegom and P. W. Wilson (eds.): *Exploring Research Frontiers in Contemporary Statistics and Econometrics*. Springer-Verlag: Berlin, pp. 65–81.

Daouia, A. and I. Gijbels (2011b), 'Robustness and inference in nonparametric partial frontier modeling'. *Journal of Econometrics* **161**, 147–165.

Daouia, A. and A. Ruiz-Gazen (2006), 'Robust nonparametric frontier estimators: Qualitative robustness and influence function'. *Statistica Sinica* **16**, 1233–1253.

Daouia, A. and L. Simar (2005), 'Robust nonparametric estimators of monotone boundaries'. *Journal of Multivariate Analysis* **96**, 311–331.

Daouia, A. and L. Simar (2007), 'Nonparametric efficiency analysis: A multivariate conditional quantile approach'. *Journal of Econometrics* **140**, 375–400.

Daraio, C. and L. Simar (2005), 'Introducing environmental variables in nonparametric frontier models: A probabilistic approach'. *Journal of Productivity Analysis* **24**, 93–121.

Daraio, C. and L. Simar (2007a), *Advanced Robust and Nonparametric Methods in Efficiency Analysis*. New York: Springer Science+Business Media, LLC.

Daraio, C. and L. Simar (2007b), 'Conditional nonparametric frontier models for convex and nonconvex technologies: A unifying approach'. *Journal of Productivity Analysis* **28**, 13–32.

Debreu, G. (1951), 'The coefficient of resource utilization'. *Econometrica* **19**, 273–292.

Deprins, D., L. Simar, and H. Tulkens (1984), 'Measuring labor inefficiency in post offices'. In: M. M. P. Pestieau and H. Tulkens (eds.): *The Performance of Public Enterprises: Concepts and Measurements*. Amsterdam: North-Holland, pp. 243–267.

Efron, B. (1979), 'Bootstrap methods: Another look at the jackknife'. *Annals of Statistics* **7**, 1–16.

Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia: Society for Industrial and Applied Mathematics. CBMS-NSF Regional Conference Series in Applied Mathematics, #38.

Efron, B. and R. J. Tibshirani (1993), *An Introduction to the Bootstrap*. London: Chapman and Hall.

Fan, Y., Q. Li, and A. Weersink (1996), 'Semiparametric estimation of stochastic production frontier models'. *Journal of Business and Economic Statistics* **14**, 460–468.

Färe, R. (1988), *Fundamentals of Production Theory*. Berlin: Springer-Verlag.

Färe, R., S. Grosskopf, and C. A. K. Lovell (1985), *The Measurement of Efficiency of Production*. Boston: Kluwer-Nijhoff Publishing.

Färe, R., S. Grosskopf, and D. Margaritis (2008), 'Productivity and efficiency: Malmquist and more'. In: H. Fried, C. A. K. Lovell, and S. Schmidt (eds.): *The Measurement of Productive Efficiency*, chapter 5. Oxford: Oxford University Press, 2nd edition, pp. 522–621.

Farrell, J. M. (1957), 'The measurement of productive efficiency'. *Journal of the Royal Statistical Society A* **120**, 253–281.

Florens, P. J. and L. Simar (2005), 'Parametric approximations of nonparametric frontiers'. *Journal of Economietrics* **124**, 91–116.

Gijbels, I., E. Mammen, B. U. Park, and L. Simar (1999), 'On estimation of monotone and concave frontier functions'. *Journal of the American Statistical Association* **94**, 220–228.

Hadley, G. (1962), *Linear Programming*. Reading, Pennsylvania: Addison-Wesley, Inc.

Hall, P. (1992), *The Bootstrap and Edgeworth Expansion*. New York: Springer-Verlag.

Hall, P. and L. Simar (2002), 'Estimating a changepoint, boundary or frontier in the presence of observation error'. *Journal of the American Statistical Association* **97**, 523–534.

Hsu, L. P. and H. Robbins (1947), 'Complete convergence and the law of large numbers'. *Proceedings of the National Academy of Sciences of the United States of America* **33**, 25–31.

Jeong, O. S., B. U. Park, and L. Simar (2010), 'Nonparametric conditional efficiency measures: asymptotic properties'. *Annals of Operational Research* **173**, 105–122.

Jeong, O. S. and L. Simar (2006), 'Linearly interpolated FDH efficiency score for nonconvex frontiers'. *Journal of Multivariate Analysis* **97**, 2141–2161.

Johnson, R. C. (1974a), 'The Hadamard product of a and a^*'. *Pacific Journal of Mathematics* **51**, 477–481.

Johnson, R. C. (1974b), 'Hadamard products of matrices'. *Linear and Multilinear Algebra* **1**, 295–307.

Jondrow, J., C. A. K. Lovell, I. S. Materov, and P. Schmidt (1982), 'On the estimation of technical inefficiency in the stochastic frontier production model'. *Journal of Econometrics* **19**, 233–238.

Kneip, A., B. Park, and L. Simar (1998), 'A note on the convergence of nonparametric DEA efficiency measures'. *Econometric Theory* **14**, 783–793.

Kneip, A., L. Simar, and I. Van Keilegom (2012a), 'Boundary estimation in the presence of measurement error with unknown variance'. Discussion paper #2012/02, Institut de Statistique Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Kneip, A., L. Simar, and P. W. Wilson (2008), 'Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models'. *Econometric Theory* **24**, 1663–1697.

Kneip, A., L. Simar, and P. W. Wilson (2011), 'A computationally efficient, consistent bootstrap for inference with non-parametric DEA estimators'. *Computational Economics* **38**, 483–515.

Kneip, A., L. Simar, and P. W. Wilson (2012b), 'Central limit theorems for DEA scores: When bias can kill the variance'. Discussion paper #2012/xx, Institut de Statistique Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Koenker, R. and G. Bassett (1978), 'Regression quantiles'. *Econometrica* **46**, 33–50.

Koopmans, C. T. (1951), 'An analysis of production as an efficient combination of activities'. In: T. C. Koopmans (ed.): *Activity Analysis of Production and Allocation*. New York: John-Wiley and Sons, Inc., pp. 33–97. Cowles Commission for Research in Economics, Monograph 13.

Korostelev, A., L. Simar, and A. B. Tsybakov (1995a), 'Efficient estimation of monotone boundaries'. *The Annals of Statistics* **23**, 476–489.

Korostelev, A., L. Simar, and A. B. Tsybakov (1995b), 'On estimation of monotone and convex boundaries'. *Publications de l'Institut de Statistique de l'Université de Paris XXXIX* **1**, 3–18.

Kumbhakar, C. S., B. U. Park, L. Simar, and E. G. Tsionas (2007), 'Nonparametric stochastic frontiers: A local likelihood approach'. *Journal of Econometrics* **137**, 1–27.

Kuosmanen, T. and M. Kortelainen (2012), 'Stochastic non-smooth envelopment of data: Semi-parametric frontier estimation subject to shape constraints'. *Journal of Productivity Analysis* **38**, 11–28.

Marcus, M. and N. Kahn (1959), 'A note on the Hadamard product'. *Canadian Mathematical Bulletin* **2**, 81–93.

Marcus, M. and R. C. Thompson (1963), 'The field of values of the Hadamard product'. *Archiv der Mathematik* **14**, 283–288.

Meeusen, W. and J. van den Broeck (1977), 'Efficiency estimation from Cobb-Douglas production functions with composed error'. *International Economic Review* **18**, 435–444.

Olesen, B. O., N. C. Petersen, and C. A. K. Lovell (1996), 'Summary of the workshop discussion'. *Journal of Productivity Analysis* **7**, 341–345.

Park, U. B., S.-O. Jeong, and L. Simar (2010), 'Asymptotic distribution of conical-hull estimators of directional edges'. *Annals of Statistics* **38**, 1320–1340.

Park, U. B., L. Simar, and C. Weiner (2000), 'FDH efficiency scores from a stochastic point of view'. *Econometric Theory* **16**, 855–877.

Park, U. B., L. Simar, and V. Zelenyuk (2008), 'Local likelihood estimation of truncated regression and its partial derivative: Theory and application'. *Journal of Econometrics* **146**, 185–2008.

Politis, N. D. and J. P. Romano (1994), 'Large sample confidence regions based on subsamples under minimal assumtions'. *Annals of Statistics* **22**, 2031–2050.

Politis, N. D. and J. P. Romano (1999), *Subsampling*. New York: Springer-Verlag, Inc.

Politis, N. D., J. P. Romano, and M. Wolf (2001), 'On the asymptotic theory of subsampling'. *Statistica Sinica* **11**, 1105–1124.

Porembski, M., K. Breitenstein, and P. Alpar (2005), 'Visualizing efficiency and reference relations in data envelopment analysis with an application to the branches of a German bank'. *Journal of Productivity Analysis* **23**, 203–221.

Racine, J. and Q. Li (2004), 'Nonparametric estimation of regression functions with both categorical and continuous data'. *Journal of Econometrics* **119**, 99–130.

Ritter, C. and L. Simar (1997), 'Pitfalls of normal–gamma stochastic frontier models'. *Journal of Productivity Analysis* **8**, 167–182.

Robinson, M. P. (1988), 'Root-n-consistent semiparametric regression'. *Econometrica* **56**, 931–954.

Schuster, F. E. (1985), 'Incorporating support constraints into nonparametric estimators of densities'. *Communications in Statistics–Theory and Methods* **14**, 1123–1136.

Scott, D. (1992), *Multivariate Density Estimation: Theory, Practice, and Visualization*. New York: John Wiley & Sons, Inc.

Serfling, J. R. (1980), *Approximation Theorems of Mathematical Statistics*. New York: John Wiley & Sons, Inc.

Sheather, J. S. and M. C. Jones (1991), 'A reliable data-based bandwidth selection method for kernel density estimation'. *Journal of the Royal Statistical Society, Series B* **53**, 683–690.

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

Silverman, W. B. (1978), 'Choosing the window width when estimating a density'. *Biometrika* **65**, 1–11.

Silverman, W. B. (1986), *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

Simar, L. (1996), 'Aspects of statistical analysis in DEA-type frontier models'. *Journal of Productivity Analysis* **7**, 177–185.

Simar, L. (2003), 'Detecting outliers in frontier models: A simple approach'. *Journal of Productivity Analysis* **20**, 391–424.

Simar, L. (2007), 'How to improve the performances of DEA/FDH estimators in the presence of noise'. *Journal of Productivity Analysis* **28**, 183–201.

Simar, L. and A. Vanhems (2012), 'Probabilistic characterization of directional distances and their robust versions'. *Journal of Econometrics* **166**, 342–354.

Simar, L., A. Vanhems, and P. W. Wilson (2012), 'Statistical inference for dea estimators of directional distances'. *European Journal of Operational Research* **220**, 853–864.

Simar, L. and P. W. Wilson (1998), 'Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models'. *Management Science* **44**, 49–61.

Simar, L. and P. W. Wilson (1999a), 'Estimating and bootstrapping Malmquist indices'. *European Journal of Operational Research* **115**, 459–471.

Simar, L. and P. W. Wilson (1999b), 'Of course we can bootstrap DEA scores! but does it mean anything? logic trumps wishful thinking'. *Journal of Productivity Analysis* **11**, 93–97.

Simar, L. and P. W. Wilson (1999c), 'Some problems with the Ferrier/Hirschberg bootstrap idea'. *Journal of Productivity Analysis* **11**, 67–80.

Simar, L. and P. W. Wilson (2000a), 'A general methodology for bootstrapping in non-parametric frontier models'. *Journal of Applied Statistics* **27**, 779–802.

Simar, L. and P. W. Wilson (2000b), 'Statistical inference in nonparametric frontier models: The state of the art'. *Journal of Productivity Analysis* **13**, 49–78.

Simar, L. and P. W. Wilson (2001a), 'Nonparametric tests of returns to scale'. *European Journal of Operational Research* **139**, 115–132.

Simar, L. and P. W. Wilson (2001b), 'Testing restrictions in nonparametric efficiency models'. *Communications in Statistics* **30**, 159–184.

Simar, L. and P. W. Wilson (2007), 'Estimation and inference in two-stage, semi-parametric models of productive efficiency'. *Journal of Econometrics* **136**, 31–64.

Simar, L. and P. W. Wilson (2008), 'Statistical inference in nonparametric frontier models: Recent developments and perspectives'. In: H. Fried, C. A. K. Lovell, and S. Schmidt (eds.): *The Measurement of Productive Efficiency*, chapter 4. Oxford: Oxford University Press, 2nd edition, pp. 421–521.

Simar, L. and P. W. Wilson (2010), 'Estimation and inference in cross-sectional, stochastic frontier models'. *Econometric Reviews* **29**, 62–98.

Simar, L. and P. W. Wilson (2011a), 'Inference by the m out of n bootstrap in nonparametric frontier models'. *Journal of Productivity Analysis* **36**, 33–53.

Simar, L. and P. W. Wilson (2011b), 'Two-Stage DEA: Caveat emptor'. *Journal of Productivity Analysis* **36**, 205–218.

Simar, L. and V. Zelenyuk (2011), 'Stochastic FDH/DEA estimators for frontier analysis'. *Journal of Productivity Analysis* **36**, 1–20.

Swanepoel, J. W. H. (1986), 'A note on proving that the (modified) bootstrap works'. *Communications in Statistics: Theory and Methods* **15**, 3193–3203.

Wheelock, C. D. and P. W. Wilson (2003), 'Robust nonparametric estimation of efficiency and technical change in U. S. commercial banking'. Unpublished working paper, The John E. Walker Department of Economics, 222 Sirrine Hall, Clemson University, Clemson, South Carolina 29634, USA.

Wheelock, C. D. and P. W. Wilson (2008), 'Non-parametric, unconditional quantile estimation for efficiency analysis with an application to Federal Reserve check processing operations'. *Journal of Econometrics* **145**, 209–225.

Wheelock, C. D. and P. W. Wilson (2011), 'Are credit unions too small?'. *Review of Economics and Statistics* **93**, 1343–1359.

Wilson, W. P. (1993), 'Detecting outliers in deterministic nonparametric frontier models with multiple outputs'. *Journal of Business and Economic Statistics* **11**, 319–323.

Wilson, W. P. (1995), 'Detecting influential observations in data envelopment analysis'. *Journal of Productivity Analysis* **6**, 27–45.

Wilson, W. P. (2004), 'A preliminary non-parametric analysis of public education and health expenditures developing countries'. Unpublished working paper, The John E. Walker Department of Economics, 222 Sirrine Hall, Clemson University, Clemson, South Carolina 29634, USA.

Wilson, W. P. (2008), 'FEAR: A software package for frontier efficiency analysis with R'. *Socio-Economic Planning Sciences* **42**, 247–254.

Wilson, W. P. (2011), 'Asymptotic properties of some non-parametric hyperbolic efficiency estimators'. In: I. Van Keilegom and P. W. Wilson (eds.): *Exploring Research Frontiers in Contemporary Statistics and Econometrics*. Berlin: Springer-Verlag, pp. 115–150.

Wood, S. F. (1973), 'The use of individual effects and residuals in fitting equaitons to data'. *Technometrics* **15**, 677–694.