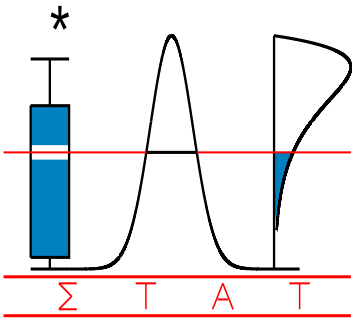
T E C H N I C A L
R E P O R T

11037

Statistical Inference for DEA Estimators of Directional Distances

SIMAR, L., VANHEMS, A. and P.W. WILSON



I A P S T A T I S T I C S
N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Statistical Inference for DEA Estimators of Directional Distances

Léopold Simar[§] Anne Vanhems[**] Paul W. Wilson[‡‡]

September 2011

Abstract

In productivity and efficiency analysis, the technical efficiency of a production unit is measured through its distance to the efficient frontier of the production set. The most familiar non-parametric methods use Farrell-Debreu, Shephard, or hyperbolic radial measures. These approaches require that inputs and outputs be non-negative, which can be problematic when using financial data. Recently, Chambers et al. (1998) have introduced directional distance functions which can be viewed as additive (rather than multiplicative) measures efficiency. Directional distance functions are not restricted to non-negative input and output quantities; in addition, the traditional input and output-oriented measures are nested as special cases of directional distance functions. Consequently, directional distances provide greater flexibility. However, until now, only FDH estimators (and their conditional and robust extensions) of directional distances have known statistical properties (Simar and Vanhems, 2010). This paper develops the statistical properties of directional DEA estimators, which are especially useful when the production set is assumed convex. We first establish that the directional DEA estimators share the known properties of the traditional radial DEA estimators. We then use these properties to develop consistent bootstrap procedures for statistical inference about directional distance, estimation of confidence intervals, and bias correction. The methods are illustrated in some empirical examples.

Keywords: Productivity, efficiency, directional distances, non-parametric frontier estimation, bootstrap, data envelopment analysis.

***JEL* Classification:** Primary C13; secondary C14

[§]Institute of Statistics, Université Catholique de Louvain-la-Neuve, Belgium. Support from the "Interuniversity Attraction Pole", Phase VI (No. P6/03) of the Belgian Science Policy and from the INRA-GREMAQ, Toulouse, France is gratefully acknowledged. Email: leopold.simar@uclouvain.be.

[**]University of Toulouse, Toulouse Business School and Toulouse School of Economics, France. Email: a.vanhems@esc-toulouse.fr.

[‡‡]The John E. Walker Department of Economics, Clemson University, South Carolina 29634, USA. Email: pww@clemson.edu.

1 Introduction

Non-parametric estimators of technical efficiency introduced by Farrell (1957) and popularized by Charnes et al. (1978) and Banker et al. (1984) based on minimal convex sets that envelop a set of observations on input and output quantities have been widely applied; Gattoufi et al. (2004) list more than 1,800 published examples. These estimators are collectively known as data envelopment analysis (DEA) estimators. *Technical efficiency* refers to whether, and to what degree, a production unit might (i) increase its output without increasing input usage, (ii) reduce its input usage without reducing its level of outputs, or (iii) simultaneously reduce input usage while increasing output quantities produced.

Until recently, technical efficiency has been measured almost exclusively in terms of either (i) or (ii), i.e., in either an input or output orientation, respectively. As discussed below in Section 2, the statistical properties of input and output oriented DEA estimators are known, and inference-making methods are available for these estimators. Färe et al. (1985) proposed measuring technical efficiency along a hyperbolic path; this approach amounts to describing technical inefficiency in terms of the maximum, feasible, multiplicative factor by which input quantities can be divided and by which output quantities can be simultaneously multiplied. The statistical properties of DEA estimators of this measure are also known, and inference methods are available, again as discussed below in Section 2. Chambers et al. (1998) proposed an additive measure of technical efficiency which allows one to consider how much might be feasibly added to a unit's output quantities and simultaneously subtracted from its input quantities; this measure is known as the *directional distance function*.

Since the appearance of Chambers et al. (1998), a number of subsequent papers using directional distance functions have appeared; examples include Fukuyama (2003), Silva-Portela et al. (2004), Ricazo-Tadeo et al. (2005), Färe et al. (2005), Park and Weber (2006a, 2006b), and many more. Directional distance functions are often used in applications involving undesirable outputs such as pollution, or in cases where observed inputs or outputs can be negative. Färe et al. (2008) discuss the importance of the directional distance function and its relation to profit efficiency, duality theory, and Luenberger, Bennet-Bowley and other indices of productivity. In addition, Briec et al. (2004) and Briec and Kerstens (2009) use directional distance functions to evaluate the performances of mutual fund managers.

Although Simar and Vanhems (2010) have derived statistical properties of the FDH estimator of directional distances, most researchers estimating directional distances have used DEA estimators. Yet, the properties of these estimators remain unknown, and consequently statistical inference has not been possible in studies using DEA estimators of directional distance functions.

This paper extends the results of Kneip et al. (2008, 2011) to DEA estimators of directional distance functions. In so doing, we provide rates of convergence as well as the limiting distribution under variable returns to scale. As will be seen, the asymptotic properties of DEA estimators of directional distance functions are similar to those of DEA estimators of input- and output-oriented distance function estimators. Although the limiting distributions derived below are not useful in a direct, practical sense for inference, the results allow us to establish consistency of sub-sampling methods along the lines of Simar and Wilson (2011) for inference, and permit DEA estimators of directional distance functions to be used for testing general hypotheses about the structure of the frontier or technology (e.g., convexity, returns to scale, etc.) along the lines of Simar and Wilson (2011). Directional distance functions allow for negative observations on inputs or outputs, and so are useful where these might occur, as in one of our empirical examples given below. Given the increasing interest in the use of DEA estimators of directional distance functions, we anticipate that our results will be useful for an increasing number of empirical researchers who, until now, have had no method for making inference or testing hypotheses when using DEA estimators of directional distances.

The paper proceeds as follows. The next section gives a careful description of the directional distance function, its relation to other distance functions, and a brief review of existing results on non-parametric distance function estimators. Section 3 introduces the non-parametric estimators (both DEA and the free-disposal hull (FDH) estimator described by Deprins et al., 1984) of directional distance functions that will be used in our empirical illustrations. The main contribution of the paper comes in Section 4, where we show that the DEA estimator of directional distances share the same properties as the DEA estimators of input and output distance functions. These results are then used in Section 5 to develop bootstrap algorithms for performing consistent inference in a practical, implementable way. In Section 6 the methods are illustrated by estimating confidence intervals for observations

on schools in the data used by Charnes et al. (1981) and for observations in a small sample of risk-returns data for Mutual Funds. Summary and conclusions are given in Section 7.

2 Radial and Directional Distance Functions

In productivity and efficiency analysis, the technical efficiency of a production unit is measured through its distance to the efficient frontier, or boundary, of the production set. Consider a set of p input quantities used to produce q output quantities. Then the production set Ψ is the set of technically feasible combinations of inputs and outputs, i.e.,

$$\Psi = \left\{(x, y) \in \mathbb{R}_+^{p+q} \mid x \text{ can produce } y\right\}. \tag{2.1}$$

The efficient frontier or boundary of Ψ, also known as the technology, is given by

$$\Psi^\partial = \left\{(x, y) \in \Psi \mid \left(\gamma^{-1}x, \gamma y\right) \notin \Psi \; \forall \; \gamma > 1\right\}. \tag{2.2}$$

The technical efficiency of a given point (x, y) is then determined by the distance to the frontier Ψ^∂. The Farrell-Debreu and their reciprocal Shephard distances (Debreu, 1951; Farrell, 1957; Shephard, 1970) are the most widely used measures of technical efficiency; both rely on multiplicative radial measures of distance. For example, the Shephard (1970) input distance function

$$\theta(x, y \mid \Psi) = \sup\left\{\theta \mid (\theta^{-1}x, y) \in \Psi\right\} \tag{2.3}$$

gives the maximum, feasible reduction in input quantities, holding output quantities fixed, for a unit operating at $(x, y) \in \Psi$. The input distance function is multiplicative in the sense that for $(x, y) \in \Psi$, we have $(\theta^{-1}x, y) \in \Psi^\partial$. Similarly, the Shephard (1970) output distance function

$$\lambda(x, y \mid \Psi) = \inf\left\{\lambda \mid (x, \lambda^{-1}y) \in \Psi\right\} \tag{2.4}$$

gives the maximum, feasible increase in output quantities, holding input quantities fixed, for a unit operating at $(x, y) \in \Psi$. The output distance function is also multiplicative in the sense that for $(x, y) \in \Psi$, we have $(x, \lambda^{-1}y) \in \Psi^\partial$. Alternatively, the hyperbolic distance function

$$\gamma(x, y \mid \Psi) = \sup\left\{\gamma \mid (\gamma^{-1}x, \gamma y) \in \Psi\right\} \tag{2.5}$$

proposed by Färe et al. (1985) and examined by Wheelock and Wilson (2008) and Wilson (2011) is also multiplicative; $\gamma(x, y \mid \Psi)$ gives the maximum feasible, equiproportionate, simultaneous reduction in input quantities and increase in output quantities.[1]

The input- and output-oriented measures of efficiency in (2.3)–(2.4) are *radial* in the sense that efficiency of a particular point is defined in terms of the feasible, equiproportionate reduction of all inputs (or increase of all outputs), holding output (or input) quantities fixed. The hyperbolic measure defined in (2.5) is also said to be radial; with $\gamma(x, y \mid \Psi)$, efficiency of a particular point is defined in terms of the feasible, equiproportionate reduction of all inputs and simultaneous feasible, equiproportionate increase of all outputs. In addition, the multiplicative construction of the measures in (2.3)–(2.5) ensures that the distance functions are independent of units of measure for either input quantities or output quantities; i.e., $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$, and $\gamma(x, y \mid \Psi)$ are homogeneous of degree zero in both input and output quantities.

Chambers et al. (1998) introduce the directional distance function

$$\delta(x, y \mid d_x, d_y, \Psi) = \sup \{\delta \mid (x - \delta d_x, y + \delta d_y) \in \Psi\}, \tag{2.6}$$

which projects the input-output vector $(x, y) \in \mathbb{R}^{p+q}$ onto the technology frontier in a direction determined by a vector $d = (-d_x, d_y)$, where $(d_x, d_y) \in \mathbb{R}^{p+q}_+$. Directional distance functions are discussed by Färe and Grosskopf (2000) and Färe et al. (2008). By construction, $\delta(x, y \mid d_x, d_y, \Psi) \geq 0$ if and only if $(x, y) \in \Psi$, and $\delta(x, y \mid d_x, d_y, \Psi) = 0$ for all $(x, y) \in \Psi^{\partial}$. In addition, the directional distance function in (2.6) nests both the input- and output-oriented measures in (2.3)–(2.4) (but not the hyperbolic distance function in (2.5)) as special cases. It is easy to show that for $d_x = x$, $d_y = 0$, and $(x, y) \in \Psi \subset \mathbb{R}^{p+q}_+$,

$$\delta(x, y \mid d_x, d_y, \Psi) = 1 - \theta(x, y \mid \Psi)^{-1}. \tag{2.7}$$

Similarly, for $d_x = 0$, $d_y = y$, and $(x, y) \in \Psi \subset \mathbb{R}^{p+q}_+$,

$$\delta(x, y \mid d_x, d_y, \Psi) = \lambda(x, y \mid \Psi)^{-1} - 1. \tag{2.8}$$

Hence, the directional distance function is more general than either the input or output distance functions.

[1] The hyperbolic measure of efficiency proposed by Färe et al. (1985) is the reciprocal of the measure that appears in (2.5).

The directional distance function is additive, as opposed to multiplicative; $\delta(x, y \mid d_x, d_y, \Psi)$ gives the amount that can be subtracted from input quantities x and simultaneously added to output quantities y in the direction $(-d_x, d_y)$ to reach the frontier. This differs from the traditional input- and output-oriented distance functions, as well as the hyperbolic distance function, used to measure efficiency.

Due to its additive (as opposed to multiplicative) nature, the directional distance function defined in (2.6) satisfies a translation property in the sense that

$$\delta(x - \eta d_x, y + \eta d_y \mid d_x, d_y, \Psi) = \delta(x, y \mid d_x, d_y, \Psi) - \eta \; \forall \; \eta \in \mathbb{R}, \tag{2.9}$$

i.e., multiplying the direction vectors by a constant η is equivalent to subtracting η from the distance function with the original direction vectors. In addition, Färe et al. (2008, p. 534) state that the directional distance function is independent of unit of measurement in the sense that

$$\delta(\beta_x \circ x, \beta_y \circ y \mid \beta_x \circ d_x, \beta_y \circ d_y, \Psi) = \delta(x, y \mid d_x, d_y, \Psi), \tag{2.10}$$

where $\beta_x \in \mathbb{R}^p_+$, $\beta_y \in \mathbb{R}^q_+$, and $\circ$ denotes the Hadamard product.[2] However, while (2.10) is true, it also indicates that if units of 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 $d_x = 1,\ d_y = 1$ or to optimize u and v to minimize distance to the (estimated) frontier. But, if one specifies $d_x = 1,\ d_y = 1$, and then changes the units of measurement, this will require re-scaling also d_x and d_y so that their no longer equal unity in order to avoid changing the value of the distance function. Hence the choice of $(1, 1)$ for (d_x, d_y) is arbitrary and somewhat meaningless. Moreover, if

[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 $c_{ij} = a_{ij}b_{ij}$; e.g., see Marcus and Kahn (1959).

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

As noted above, the relations in (2.7)–(2.8) require $\Psi \subset \mathbb{R}^{p+q}_+$. In certain situations, however, it is conceivable that the production process might be such that negative inputs or outputs are possible, i.e., if $\Psi \cap \mathbb{R}^{p+q}_+ \neq \Psi$. In such cases, the input and output distance functions are undefined, but the directional distance function remains defined due to its additive (instead of multiplicative) construction. If, for example, $y < 0$, $d_x = 0$, and $d_y = -y$ (so that $d_y > 0$), then there exists a value $\delta > 1$ such that $(x, y + \delta d_y) = (x, (1 - \delta)y) \in \Psi^\partial$. The relation between the directional distance function and the output distance function in (2.8) has no meaning in this case, but the directional distance function remains well-defined.

Non-parametric estimation of technical efficiency typically involves estimating the production set Ψ using a sample of observations $\mathcal{X}_n = \{(X_i, Y_i)\}_{i=1}^n$. The standard non-parametric estimators of Ψ are (i) the Free Disposal Hull (FDH) of the sample observations, proposed by Deprins et al. (1984), and (ii) the convex hull of the FDH of the sample observations, proposed by Farrell (1957) and popularized by Charnes et al. (1978). The latter approach is commonly referred to as Data Envelopment Analysis (DEA). The FDH approach requires an assumption of free disposability of the inputs and the outputs, while the DEA approach requires the additional assumption that the production set Ψ is convex.

Properties of FDH and DEA estimators of the radial distance functions defined in (2.3)–(2.5) are now well-known. Park et al. (2000) established asymptotic results for the input and output oriented FDH estimators, and Kneip et al. (1998) and Kneip et al. (2008) derived asymptotic results for the input and output oriented DEA estimators under variable returns to scale (VRS). Wilson (2011) extended these results to FDH and DEA estimators of the hyperbolic distance function defined in (2.5), while Park et al. (2010) extended the DEA results to cases where returns to scale are constant. To briefly summarize, under mild regularity conditions, the FDH and DEA estimators of the radial distance functions, when scaled by appropriate rates of convergence, have non-degenerate limiting distributions (more details are given below).[3] However, each of the limiting distributions depends on unknown parameters, which in turn depend on characteristics of the Data Generating Process (DGP).

[3] By "non-degenerate" we mean that the distribution is not a probability mass concentrated at a single point.

The limiting distributions of the FDH estimators can be written in closed form, but it is difficult to estimate the unknown parameters. The limiting distributions of the DEA estimators have no closed, analytical expression; in principle, the distributions could be estimated using simulation methods, but it is not clear how the unknown parameters might be estimated.

Bootstrap methods provide the only practical approach for inference using either FDH or DEA estimators of the input, output, or hyperbolic distance functions. Jeong and Simar (2006) prove that sub-sampling provides a consistent bootstrap approximation for FDH estimators of the radial distance functions. Kneip et al. (2008) prove consistency of two different bootstrap methods for DEA estimators of radial distance functions; the first employs a complicated, double-smoothing technique, while the second uses sub-sampling. Kneip et al. (2011) propose a simplified version of the double-smooth bootstrap and prove its consistency; this approach is useful for estimating confidence intervals for the efficiency of individual points in the input-output space, but is not useful for testing hypotheses about the frontier, etc. For purposes of hypothesis testing (e.g., testing hypotheses about returns to scale, separability, convexity of Ψ, etc.), the sub-sampling techniques proposed by Simar and Wilson (2011) offer the only practical approach available to date. The methods developed by Simar and Wilson (2011) have been used for testing returns to scale or testing convexity in Simar and Wilson (2011), testing restrictions in Schubert and Simar (2011), testing separability in Daraio et al. (2010), etc.

Simar and Vanhems (2010) provide asymptotic results for FDH estimators of directional distances, as well as results for robust order-m and order-α quantile estimators of directional distances; their results also allow (in the FDH case) conditioning on environmental variables. The results reveal that the FDH estimators of directional distances possess asymptotic properties similar to those of the FDH estimators of radial distance functions. To date, however, no such results exist for DEA estimators of directional distances. This paper provides the missing piece in the collection of results on asymptotics of non-parametric distance function estimators by developing asymptotic results for DEA estimators of directional distance functions. We next describe the estimators; the asymptotic results follow.

3 Non-parametric Estimation of Directional Distances

3.1 The Model

For the remainder of the paper, unless otherwise stated, let $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$ denote vectors of input and output quantities. The vectors x and y need not be non-negative. We make standard assumptions on the production set Ψ by adopting those of Shephard (1970) and Färe et al. (1985).

Assumption 3.1. *Ψ is convex.*

Assumption 3.2. *Ψ is closed.*

Assumption 3.3. *Both inputs and outputs are strongly disposable; i.e., for $\widetilde{x} \geq x$, $\widetilde{y} \leq y$, if $(x, y) \in \Psi$ then $(\widetilde{x}, y) \in \Psi$ and $(x, \widetilde{y}) \in \Psi$.*[4]

Convexity of Ψ is required for consistency of DEA estimators, which estimate Ψ by convex sets. Assumption 3.3 amounts to assuming monotonicity of isoquants (both input and output), and is a common assumption in production analysis. Let the operator $\mathcal{F}(\cdot)$ denote the free-disposal hull of a set in $\mathbb{R}^{p+q}$. Then for any set $\mathcal{A} \subseteq \mathbb{R}^{p+q}$,

$$\mathcal{F}(\mathcal{A}) = \bigcup_{(x,y)\in\mathcal{A}} \left\{(\widetilde{x}, \widetilde{y}) \in \mathbb{R}^{p+q} \mid \widetilde{y} \leq y,\ \widetilde{x} \geq x\right\}. \tag{3.1}$$

Now let $\mathcal{C}(\Psi)$ denote the convex hull of $\mathcal{F}(\Psi)$. In some cases, it is also useful to consider the conical hull of $\mathcal{F}(\Psi)$, which we denote by $\mathcal{V}(\Psi)$. If Ψ^∂ exhibits constant returns to scale (CRS) everywhere, then $\Psi = \mathcal{V}(\Psi)$. Assumptions 3.2–3.3 ensure that $\Psi = \mathcal{F}(\Psi) \subseteq \mathcal{C}(\Psi)$, while Assumption 3.1 implies $\Psi = \mathcal{F}(\Psi) = \mathcal{C}(\Psi)$. The notation introduced here will be useful for describing the test of convexity that we introduce later.

When estimating Ψ or Ψ^∂ from the sample $\mathcal{X}_n = \{(X_i, Y_i)\}_{i=1}^n$, additional assumptions are needed to complete the statistical model. Following Kneip et al. (1998), Kneip et al. (2008, 2011), and Park et al. (2000), we assume the following.

Assumption 3.4. *The n observations in $\mathcal{X}_n$ are identically, independently distributed (iid) random variables on the attainable set Ψ.*

[4] Note that as usual, inequalities involving vectors are defined on an element-by-element basis.

Assumption 3.5. *The random variables (X, Y) possess a joint density f with compact support $\mathcal{D} \subset \Psi$; f is continuous on $\mathcal{D}$; and $f(x, y) > 0 \; \forall \; (x, y) \in \Psi^{\partial}$.*

Assumption 3.6. *The function $\delta(x, y \mid d_x, d_y, \Psi)$ is twice continuously differentiable for all $(x, y) \in \Psi$ and for all $(d_x, d_y)' \in \mathbb{R}_+^{p+q}$.*

Assumption 3.5 guarantees a positive probability of observing production units close to the boundary Ψ^{∂} when the sample size increases. Assumption 3.6 imposes some smoothness on the boundary of Ψ; the condition given here is sufficient, but stronger than necessary for establishing consistency of the DEA distance function estimators. A weaker, but more complicated, assumption was used in Kneip et al. (1998).

Deprins et al. (1984) proposed estimating Ψ by the free-disposal hull of the sample observations in $\mathcal{X}_n$, i.e., by

$$\widehat{\Psi}_{\text{FDH}}(\mathcal{X}_n) = \mathcal{F}(\mathcal{X}_n) = \bigcup_{(x_i, y_i) \in \mathcal{X}_n} \left\{ (x, y) \in \mathbb{R}^{p+q} \mid y \leq Y_i, \; x \geq X_i \right\}. \tag{3.2}$$

FDH estimators $\widehat{\theta}_{\text{FDH}}(x, y \mid \mathcal{X}_n)$, $\widehat{\lambda}_{\text{FDH}}(x, y \mid \mathcal{X}_n)$, $\widehat{\gamma}_{\text{FDH}}(x, y \mid \mathcal{X}_n)$, and $\widehat{\delta}_{\text{FDH}}(x, y \mid d_x, d_y, \mathcal{X}_n)$ of the distance functions defined in (2.6)–(2.5) are obtained by replacing Ψ on the right-hand sides of (2.3)–(2.6) with $\widehat{\Psi}_{\text{FDH}}(\mathcal{X}_n)$.

Under Assumptions 3.2–3.6, and for $(x, y) \in \Psi \subset \mathbb{R}_+^{p+q}$, Park et al. (2000) prove consistency and derive limiting distributions of the FDH estimators of the input and output distance functions in (2.3)–(2.4). Wilson (2011) extends these results to the FDH estimator of the hyperbolic distance function in (2.5), while Simar and Vanhems (2010) extend the results to the FDH estimator of the directional distance function in (2.6). In each case, the rate of convergence is shown to be $n^{1/(p+q)}$; moreover, each of the FDH distance function estimators is shown to have an asymptotic distribution related to the Weibull distribution. In the case of the FDH estimator of the directional distance function, the assumption that $(x, y) \in \Psi \subset \mathbb{R}_+^{p+q}$ is not needed.

The VRS-DEA estimator of Ψ under variable returns to scale is the convex hull of $\mathcal{F}(\mathcal{X}_n)$, and is given by

$$\widehat{\Psi}_{\text{VRS}}(\mathcal{X}_n) = \left\{ (x, y) \in \mathbb{R}^{p+q} \mid y \leq \boldsymbol{Y}\boldsymbol{\omega}, \; x \geq \boldsymbol{X}\boldsymbol{\omega}, \; \boldsymbol{i}_n'\boldsymbol{\omega} = 1, \; \boldsymbol{\omega} \in \mathbb{R}_+^n \right\}, \tag{3.3}$$

where $\boldsymbol{X} = \big(x_1, \; \ldots, \; x_n\big)$ and $\boldsymbol{Y} = \big(y_1, \; \ldots, \; y_n\big)$ are $(p \times n)$ and $(q \times n)$ matrices of input and output vectors, respectively; $\boldsymbol{i}_n$ is an $(n \times 1)$ vector of ones, and $\boldsymbol{\omega}$ is a $(n \times 1)$ vector of weights.

The corresponding VRS-DEA estimators $\widehat{\theta}_{\text{VRS}}(x, y \mid \mathcal{X}_n)$, $\widehat{\lambda}_{\text{VRS}}(x, y \mid \mathcal{X}_n)$, $\widehat{\gamma}_{\text{VRS}}(x, y \mid \mathcal{X}_n)$, and $\widehat{\delta}_{\text{VRS}}(x, y \mid d_x, d_y, \mathcal{X}_n)$ of the distance functions defined in (2.3)–(2.6) are obtained by replacing Ψ on the right-hand sides of (2.3)–(2.6) with $\widehat{\Psi}_{\text{VRS}}(\mathcal{X}_n)$. Kneip et al. (1998) prove that $\widehat{\theta}_{\text{VRS}}(x, y \mid \mathcal{X}_n)$ and $\widehat{\lambda}_{\text{VRS}}(x, y \mid \mathcal{X}_n)$ are consistent estimators of $\theta(x, y \mid \Psi)$ and $\lambda(x, y \mid \Psi)$ and converge at rate $n^{-2/(p+q+1)}$ under Assumptions 3.1–3.6 and the additional assumption that $(x, y) \in \Psi \subset \mathbb{R}_+^{p+q}$; limiting distributions are established by Kneip et al. (2008). Wilson (2011) extends these results by proving, under the same assumptions, that the VRS-DEA estimator $\widehat{\gamma}_{\text{VRS}}(x, y \mid \mathcal{X}_n)$ consistently estimates the hyperbolic distance function $\gamma(x, y \mid \Psi)$ defined in (2.5). The convergence rate is again $n^{-2/(p+q+1)}$ and the limiting distribution is similar to that of the input and output-oriented VRS-DEA estimators.[5]

While the limiting distribution of the VRS-DEA estimators of the input, output, and hyperbolic distance functions in (2.3)–(2.5) is known, there exists no closed, analytical form for this distribution. Consequently, bootstrap methods provide the only feasible, practical approach to inference (e.g., see Simar and Wilson, 1998, 2008, 2011 and Kneip et al., 2008, 2011). As discussed above, the radial nature of these distance functions and their estimators precludes negative values of inputs or outputs. Although directional distance functions allow for this possibility, to date, the asymptotic properties (i.e., convergence rates, limiting distributions, etc.) of VRS-DEA estimators of directional distance functions remain unknown.

The results that are available so far indicate that VRS-DEA distance function estimators have better rates of convergence than their corresponding FDH distance function estimators. However, it should be clear that if the attainable set Ψ is not convex, VRS-DEA estimators are inappropriate, providing non consistent estimators of the attainable set and consequently, statistically inconsistent estimates of any measure of distance to the frontier of Ψ. This suggests that convexity of Ψ should be tested if VRS-DEA estimators are to be used. Simar and Wilson, 2011 develop sub-sampling methods for testing the convexity assumption using

[5] The constant-returns-to-scale version of the DEA estimator of Ψ is obtained by dropping the constraint $\boldsymbol{i}_n\boldsymbol{\omega} = 1$ from the right-hand side of (3.3). In this case, Ψ is estimated by the conical hull of $\mathcal{F}(\mathcal{X}_n)$; Park et al. (2010) prove, under Assumptions 3.1–3.6 and the additional assumption that $\Psi = \mathcal{V}(\Psi)$, that the corresponding estimators of the input and output distance functions defined in (2.3)–(2.4) converge at the rate $n^{2/(p+q)}$. In addition, it is trivial to show that under globally constant returns to scale, $\gamma(x, y \mid \Psi) = \sqrt{\theta(x, y \mid \Psi)} = 1/\sqrt{\lambda(x, y \mid \Psi)}$. Consequently, under constant returns to scale, a consistent estimator of $\gamma(x, y \mid \Psi)$ is given by the square root of the input distance function estimator obtained by dropping the constraint $\boldsymbol{i}_n'\boldsymbol{\omega}$ in (3.3).

either input, output, or hyperbolic distance functions. Below, in Section 4, we derive asymptotic properties for VRS-DEA estimators of the directional distance function; this allows extension of the methods of Simar and Wilson (2011) for estimating confidence intervals and for hypothesis testing to directional distances. Before turning to the asymptotic results in Section 4, however, we first give a careful description of the FDH and VRS-DEA estimators of the directional distance function in (2.6).

3.2 Non-parametric Estimators of Directional Distances

As noted above, substituting $\widehat{\Psi}_{\text{FDH}}(\mathcal{X}_n)$ for Ψ in (2.6) yields the FDH estimator $\widehat{\delta}_{\text{FDH}}(x, y \mid d_x, d_y, \Psi)$ of $\delta(x, y \mid d_x, d_y, \Psi)$ for the point (x, y) with direction given by $d_0 = (-d_x, d_y)$. Simar and Vanhems (2010) give a probabilistic formulation of directional distances which permits straightforward derivation of the asymptotic properties of the FDH estimator of directional distances. Their main result, which will be needed below, appears in Theorem 4.1 of Simar and Vanhems (2010):

Theorem 3.1. *Under Assumptions A2–A5,*

$$n^{1/(p+q)}\left(\delta(x, y \mid d_x, d_y, \Psi) - \widehat{\delta}_{\text{FDH}}(x, y \mid d_x, d_y, \mathcal{X}_n)\right) \xrightarrow{\mathcal{L}} \text{We}\left(\mu_{(x,y)}^{p+q}, \frac{p+q}{\eta_{(x,y)}}\right), \tag{3.4}$$

for any (x, y) in the interior of Ψ, where $\text{We}(\cdot, \cdot)$ is the Weibull distribution with constants $\mu_{(x,y)} > 0$ and $\eta_{(x,y)} > 0$ completely determined by the DGP.

A proof is given in Simar and Vanhems (2010). Extensions to the robust order-m and order-α quantile directional distance and to conditional directional distance are also available in Simar and Vanhems, 2010.

Alternatively, substituting $\widehat{\Psi}_{\text{VRS}}(\mathcal{X}_n)$ for Ψ in (2.6) yields the VRS-DEA estimator

$$\widehat{\delta}_{\text{VRS}}(x, y \mid d_x, d_y, \mathcal{X}_n) = \max_{\delta, \boldsymbol{\omega}} \left\{\delta \mid y + \delta d_y \leq \boldsymbol{Y}\boldsymbol{\omega},\ x - \delta d_x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\} \tag{3.5}$$

of $\delta(x, y \mid d_x, d_y, \Psi)$ for the point (x, y) with direction given by $d_0 = (-d_x, d_y)$, where $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{\omega}$, and $\boldsymbol{i}_n$ are defined as before. The estimator $\widehat{\delta}_{\text{VRS}}(x, y \mid d_x, d_y, \mathcal{X}_n)$ measures the distance from a point (x, y) to the boundary of the convex hull of the $\mathcal{F}(\mathcal{X}_n)$ following the direction $(-d_x, d_y)$. In the next section we show that the VRS-DEA estimator in (3.5) shares the asymptotic properties of the VRS-DEA estimators of the radial input and output distance

functions, as described in Kneip et al. (2008), as well as the asymptotic properties of the VRS-DEA estimator of the hyperbolic distance function, as described in Wilson (2011).

As previously noted in footnote 5, dropping the constraint $\boldsymbol{i}_n'\boldsymbol{\omega} = 1$ in (3.3) gives the CRS-DEA estimator $\widehat{\Psi}_{\text{CRS}}(\mathcal{X}_n)$ of Ψ. This in turn leads to the CRS-DEA estimator $\widehat{\delta}_{\text{CRS}}(x, y \mid d_x, d_y, \mathcal{X}_n)$ of $\delta(x, y \mid d_x, d_y, \Psi)$; the estimator resembles the one in (3.5) except that the constraint $\boldsymbol{i}_n'\boldsymbol{\omega} = 1$ is omitted. In the next section, we focus primarily on properties of the VRS-DEA estimator of the directional distance function, followed by a brief discussion of properties of the CRS-DEA estimator.

4 Asymptotic Distribution Theory

In order to derive the statistical properties of the VRS-DEA estimator of directional distance functions, we transform the coordinate system in order to represent both the frontier and its estimator in terms of simple, scalar-valued functions. Jeong and Simar (2006) used a similar device to construct and analyze a linearly interpolated version of the FDH estimator; more recently, Kneip et al. (2008, 2011) used the same approach to derive asymptotic properties of VRS-DEA estimators of input- and output-oriented distance functions. Using the approach here, we are able to rely on results from Kneip et al. (2008, 2011) to establish asymptotic properties for VRS-DEA estimators of directional distance functions; in addition, by using a similar framework, we are able to show that existing bootstrap methods for radial efficiency measures are easy to adapt to directional measures of efficiency.

Denote the ordered pair $(x, y) \in \Psi$ by w, and let $r = p + q$ be the length of the vector w. Let $w_0 = (x_0, y_0)$ denote a specific point of interest. Suppose that we want to estimate the distance from w_0 to the frontier of Ψ in the direction $d_0 = (-d_x, d_y)$, where $d_x, d_y \geq 0$. Let $\{v_j \mid j = 1, \ldots, r-1\}$ denote an orthonormal basis for d_0, and let $\boldsymbol{V}$ be the $r \times (r-1)$ matrix whose jth column is v_j.[6]

Now consider the linear transformation from $\mathbb{R}^r$ to $\mathbb{R}^r$ given by

$$h_{w_0} \colon w \mapsto \xi = \boldsymbol{T}\,(w - w_0), \tag{4.1}$$

where $\boldsymbol{T}' = \begin{pmatrix} \boldsymbol{V} & \frac{d_0}{||d_0||} \end{pmatrix}$ is a $r \times r$ orthogonal matrix and $||d_0|| = \sqrt{d_0' d_0}$ is the Euclidean

[6] Various methods exist for computing an orthonormal basis of a vector; e.g., see Jeong and Simar (2006) or Anderson et al. (1999).

norm of d_0; then $\boldsymbol{T}'\boldsymbol{T} = \boldsymbol{I}_r$ where $\boldsymbol{I}_r$ denotes the $r \times r$ identity matrix. This one-to-one transformation can be inverted, i.e., $w = w_0 + \boldsymbol{T}'\xi$; also note that $h_{w_0}(w_0) = 0$.

To see the consequence of this transformation, partition ξ as $\xi' = \begin{pmatrix} z' & u \end{pmatrix}$, where $z = \boldsymbol{V}'(w - w_0) \in \mathbb{R}^{r-1}$ and $u = d_0'(w - w_0)||d_0||^{-1} \in \mathbb{R}$. The translation $(w - w_0)$ places the origin of the new coordinate system at the point w_0; the rotation $\boldsymbol{T}$ puts one coordinate (u) in the direction d_0 and the $r - 1$ remaining coordinates (z) are orthogonal to d_0 (and hence to the u-axis). Figure 1 illustrates the transformation for the bivariate case with $p = q = 1$.

Applying the transformation in (4.1) to each observation (X_i, Y_i) in the observed sample $\mathcal{X}_n$ results in the transformed sample $\mathcal{Z}_n = \{(Z_i, U_i)\}_{i=1}^n$, where $Z_i = \boldsymbol{V}'(\begin{bmatrix} X_i' & Y_i' \end{bmatrix}' - w_0) \in \mathbb{R}^{r-1}$ and $U_i = ||d_0||^{-1}d_0'(\begin{bmatrix} X_i' & Y_i' \end{bmatrix}' - w_0) \in \mathbb{R}$.

In the new coordinate system, the attainable set Ψ is represented by

$$\Gamma(w_0) = \{\xi \in \mathbb{R}^r \mid \xi = h_{w_0}(w),\ w \in \Psi\}. \tag{4.2}$$

This representation of Ψ depends on and is from the perspective of the point of interest w_0. The efficient frontier of Ψ can now be represented in the transformed coordinate system in terms of the scalar-valued function

$$\phi(z \mid w_0) = \sup\left\{u \mid \xi = \begin{pmatrix} z' & u \end{pmatrix}' \in \Gamma_{w_0}\right\}. \tag{4.3}$$

Figure 1 illustrates the representation of the frontier of Ψ in terms of the function $\phi(z \mid w_0)$ when $p = q = 1$. Representing the frontier in terms of a scalar-valued function permits the attainable set $\Gamma(w_0)$ to be described in terms of this function; i.e.,

$$\Gamma(w_0) = \{\xi = \begin{pmatrix} z' & u \end{pmatrix}' \in \mathbb{R}^r \mid u \leq \phi(z \mid w_0)\}. \tag{4.4}$$

It is easy to see that for a generic point $w = (x, y) \in \Psi$, and the direction vector d_0,

$$\delta(x, y \mid d_x, d_y, \Psi) = ||d_0||^{-1}(\phi(z \mid w_0) - u), \tag{4.5}$$

since $\begin{pmatrix} z' & u \end{pmatrix} = \xi'$ and $\xi = h_{w_0}(w)$. For the point of interest w_0, we have $h_{w_0}(w_0) = 0$ (since $z = 0$ and $u = 0$ at w_0), and hence

$$\delta(x, y \mid d_x, d_y, \Psi) = ||d_0||^{-1}\phi(0 \mid w_0). \tag{4.6}$$

The empirical analog of the true model replaces Ψ with its VRS-DEA estimator $\widehat{\Psi}_{\text{VRS}}(\mathcal{X}_n)$. Let $\boldsymbol{W} = \begin{pmatrix}\boldsymbol{X}' & \boldsymbol{Y}'\end{pmatrix}'$; then $\boldsymbol{W}$ is a $r \times n$ matrix containing the sample observations. The VRS-DEA estimator (3.3) can be written as

$$\widehat{\Psi}_{\text{VRS}}(\mathcal{X}_n) = \left\{ w \in \mathbb{R}^r \mid w = \boldsymbol{W}\boldsymbol{\omega} + \boldsymbol{I}_r \begin{pmatrix} \beta_x \\ -\beta_y \end{pmatrix}, \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}^n_+,\ \beta_x \in \mathbb{R}^p_+,\ \beta_y \in \mathbb{R}^q_+ \right\}. \tag{4.7}$$

The β_x and β_y are introduced in (4.7) to replace the inequalities in (3.3), and to ensure the free disposability of the estimator. In terms of the transformed coordinate system, the VRS-DEA estimator of $\Gamma(w_0)$ can be written as

$$\widehat{\Gamma}_{\text{VRS}}(\mathcal{Z}_n, w_0) = \left\{ \xi \mid \xi = \boldsymbol{T}\boldsymbol{W}\boldsymbol{\omega} + \boldsymbol{T} \begin{pmatrix} \beta_x \\ -\beta_y \end{pmatrix} - \boldsymbol{T}w_0,\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}^n_+,\ \beta_x \in \mathbb{R}^p_+,\ \beta_y \in \mathbb{R}^q_+ \right\}. \tag{4.8}$$

The boundary of $\widehat{\Psi}_{\text{VRS}}(\mathcal{X}_n)$ can now be described in the transformed coordinate system by the scalar-valued function

$$\widehat{\phi}_{\text{VRS}}(z \mid \mathcal{Z}_n, w_0) = \sup \left\{ u \mid \xi = \begin{pmatrix} z' & u \end{pmatrix}' \in \widehat{\Gamma}_{\text{VRS}}(\mathcal{Z}_n, w_0) \right\}. \tag{4.9}$$

Hence, the set $\widehat{\Gamma}_{\text{VRS}}(\mathcal{Z}_n, w_0)$ can be represented equivalently as

$$\widehat{\Gamma}_{\text{VRS}}(\mathcal{Z}_n, w_0) = \left\{ \xi = \begin{pmatrix} z' & u \end{pmatrix}' \in \mathbb{R}^r \mid u \leq \widehat{\phi}_{\text{VRS}}(z \mid \mathcal{Z}_n, w_0) \right\}. \tag{4.10}$$

In Figure 1, the VRS-DEA frontier is depicted by the dashed line, corresponding to $\widehat{\phi}_{\text{VRS}}(z \mid \mathcal{Z}_n, w_0)$. At the point w_0, we have the frontier point $\phi(0 \mid w_0)$ and its VRS-DEA estimate $\widehat{\phi}_{\text{VRS}}(0 \mid \mathcal{Z}_n, w_0)$ on the u-axis, i.e. for $z = 0$.

As a practical matter, for any point z, $\widehat{\phi}_{\text{VRS}}(z \mid \mathcal{Z}_n, w_0)$ can be obtained by a solving the simple linear program

$$\begin{aligned} \widehat{\phi}_{\text{VRS}}(z \mid \mathcal{Z}_n, w_0) = \max_{u, \boldsymbol{\omega}, \beta_x, \beta_y} \Big\{ u \mid & \boldsymbol{V}'\boldsymbol{W}\boldsymbol{\omega} + \boldsymbol{V}' \begin{pmatrix} \beta_x' & -\beta_y' \end{pmatrix}' = z + \boldsymbol{V}'w_0, \\ & u - \frac{d_0'}{||d_0||}\boldsymbol{W}\boldsymbol{\omega} - \frac{d_0'}{||d_0||} \begin{pmatrix} \beta_x' & -\beta_y' \end{pmatrix}' = -\frac{d_0'}{||d_0||}w_0, \\ & \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}^n_+,\ \beta_x \in \mathbb{R}^p_+,\ \beta_y \in \mathbb{R}^q_+ \Big\}. \end{aligned} \tag{4.11}$$

The VRS-DEA estimator of the directional distance at any point $w = (x, y) \in \Psi$, given the direction vector d_0, is

$$\widehat{\delta}_{\text{VRS}}(x, y \mid d_x, d_y, \mathcal{X}_n) = ||d_0||^{-1}(\widehat{\phi}_{\text{VRS}}(z \mid \mathcal{Z}_n, w_0) - u). \tag{4.12}$$

For the point of interest w_0, we have

$$\widehat{\delta}_{\text{VRS}}(x, y \mid d_x, d_y, \mathcal{X}_n) = ||d_0||^{-1}\widehat{\phi}_{\text{VRS}}(0 \mid \mathcal{Z}_n, w_0). \tag{4.13}$$

Careful examination of (4.6) and (4.13) reveals that we have an identical mathematical formulation to define the directional distance and its VRS-DEA estimator, as we had for the Farrell input (or output) radial distances considered in Kneip et al. (2008, 2011) (e.g., see equations (3.10) and (3.17) in Kneip et al., 2011). Consequently, under the regularity conditions described in Kneip et al. (2011) and summarized above, we obtain the following result.

Theorem 4.1. *Under the regularity conditions given by assumptions 3.1–3.6, as $n \to \infty$, for any point $(x, y) \in \Psi$, and for any direction $d_0 = (-d_x, d_y)$ where $d_x, d_y \geq 0$,*

$$n^{2/(p+q+1)}\left(\widehat{\delta}_{VRS}(x, y \mid d_x, d_y, \mathcal{X}_n) - \delta(x, y \mid d_x, d_y, \Psi)\right) \xrightarrow{\mathcal{L}} Q, \tag{4.14}$$

where Q is a random variable having a non-degenerate distribution with finite mean μ_Q and finite variance $\sigma_Q^2 > 0$.

The proof of Theorem 4.1 follows directly using the transformation introduced above and the arguments in Kneip et al. (2011); here, we give an heuristic description.

First, by Lemma 1 in Kneip et al. (2008) (where the function d_x corresponds to the function $\phi(. \mid w_0)$ in the present notation) strict convexity of Ψ implies that the function $\phi(z \mid w_0)$ is convex. Second, the smoothness of the frontier assumed in Assumption 3.6 ([*A*5]) implies that the function $\phi(z \mid w_0)$ is twice continuously differentiable for all points in a neighborhood of $z = 0$, with a positive semidefinite matrix of second derivatives at $z = 0$. By Lemma 2 in Kneip et al. (2008), the sampling distribution of $\widehat{\delta}_{\text{VRS}}(x, y \mid d_x, d_y, \mathcal{X}_n) - \delta(x, y \mid d_x, d_y, \Psi)$ can be reformulated in terms of $\phi(\cdot \mid w_0)$ and of the joint density of $\delta(X_i, Y_i \mid d_x, d_y, \Psi)$ and Z_i. Theorem 1 in Kneip et al. (2008) establishes the "local" nature of VRS-DEA estimators when Ψ is convex (as assumed in Assumption 3.1); hence the value of the directional VRS-DEA estimator $\widehat{\delta}_{\text{VRS}}(x, y \mid d_x, d_y, \mathcal{X}_n)$ is determined by observations lying in a small neighborhood of the frontier point $(x - \delta d_x,\ y + \delta d_y)$, where $\delta = \delta(x, y \mid d_x, d_y, \Psi))$. The final result is obtained by reformulating Theorem 2 of Kneip et al. (2008) to obtain the result in Theorem 4.1 above. The asymptotic properties of VRS-DEA estimators of

directional distances are identical to the asymptotic properties of VRS-DEA estimators of radial efficiency measures.

The result in Theorem 4.1 is not surprising since the usual input- and output-oriented distance functions (and their estimators) can be written as special cases of directional distance functions as shown in (2.7)–(2.8). In addition, VRS-DEA estimators of the various distance functions $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$, $\gamma(x, y \mid \Psi)$, and $\delta(x, y \mid d_x, d_y, \Psi)$ are *plug-in* estimators in the sense that the VRS-DEA estimators are obtained by replacing Ψ in the definition of a particular distance function with $\widehat{\Psi}_{\text{VRS}}$. Consequently, the statistical properties of VRS-DEA estimators of distance to the boundary of Ψ necessarily depend on the properties of the VRS-DEA estimator $\widehat{\Psi}_{\text{VRS}}$ of Ψ. The VRS-DEA distance function estimators are, in each case, smooth functions of $\widehat{\Psi}_{\text{VRS}}$, and hence share similar asymptotic properties.

In the cases where Ψ is equivalent to its conical hull, the frontier of the production set Ψ is said to exhibit globally constant returns to scale. In such cases, one can consistently estimate Ψ by $\widehat{\Psi}_{\text{CRS}}$ described above. Substituting this estimator for Ψ in (2.6) gives an estimator $\widehat{\delta}_{\text{CRS}}(x, y \mid d_x, d_y, \mathcal{X}_n)$ of $\delta(x, y \mid d_x, d_y, \Psi)$ under constant returns to scale; the estimator is similar to the VRS-DEA estimator in (3.5), except that the constraint $\boldsymbol{i}_n'\boldsymbol{\omega} = 1$ is omitted. Asymptotic results for $\widehat{\delta}_{\text{CRS}}(x, y \mid d_x, d_y, \mathcal{X}_n)$ are similar to those for $\widehat{\delta}_{\text{CRS}}(x, y \mid d_x, d_y, \mathcal{X}_n)$, and are given in the following Theorem:

Theorem 4.2. *Assume that $\Psi = \mathcal{V}(\Psi)$, and that the conditions given by Assumptions 3.1–3.6 hold. Then as $n \to \infty$, for any point $(x, y) \in \Psi$, and for any direction $d_0 = (-d_x, d_y)$ where $d_x, d_y \geq 0$,*

$$n^{2/(p+q)} \left(\widehat{\delta}_{CRS}(x, y \mid d_x, d_y, \mathcal{X}_n) - \delta(x, y \mid d_x, d_y, \Psi) \right) \xrightarrow{\mathcal{L}} Q_*, \tag{4.15}$$

where Q_* is a random variable having a non-degenerate distribution with finite mean μ_{Q_*} and finite variance $\sigma^2_{Q_*} > 0$.

We do not give a formal proof here in order to save space; however, the proof of Theorem 4.2 is straight-forward, relying on the the transformation introduced above and the reasoning in Park et al. (2010). Note that here, with constant returns to scale, the convergence rate is $n^{2/(p+q)}$; in Theorem 4.1 with variable returns to scale, the convergence rate is slower, i.e., $n^{2/(p+q+1)}$.

5 Bootstrap and Inference

The asymptotic results obtained in Section 4 are not particularly useful from a practical point of view since the limiting distribution has no closed form. The asymptotic results established in Section 4 are important, however, for establishing validity of bootstrap methods that are useful in applications for making inference about the efficiency of individual points as well as tests of general hypotheses regarding the production set Ψ (e.g., tests of convexity, returns to scale, or various restrictions that might be imposed).

Kneip et al. (2008) propose two bootstrap procedures to estimate confidence intervals for the efficiencies of individual points and prove consistency of both methods. The first approach uses sub-sampling where bootstrap samples of size $m < n$ are drawn independently, with replacement from the empirical distribution of the original sample $\mathcal{X}_n$. Simulation results, however, indicate that the coverage of bootstrap estimates of confidence intervals depends critically on the sub-sample size m. Kneip et al. were unable to offer a practical method for choosing the sub-sample size. Their second, double-smoothing approach involves bootstrap samples of size n, and requires smoothing both the joint distribution of inputs and outputs as well as the initial frontier estimate. The double-smoothing method requires solving a large number of intermediate linear programs as well as selecting values for two bandwidths, making the method computationally intensive.

For the case of confidence intervals for the input distance function defined in 2.3, Kneip et al. (2011) establish validity of a simplified, computationally efficient version of the double-smooth bootstrap developed by Kneip et al. (2008). The idea is to construct bootstrap samples by drawing "naively," i.e., from the empirical distribution of observed input-output pairs, while replacing draws of observations "near" the estimated frontier with draws from a uniform distribution. The method requires smoothing the initial frontier estimate as in the original double-smooth method of Kneip et al. (2008), but avoids the need for solving intermediate linear programs on each bootstrap replication. Consequently, the method is much faster than the original double-smoothing approach, with computational time of the same order as a simple, naive (but inconsistent) bootstrap. The simplified method also requires two bandwidths, but one (controlling the size of the neighborhood "near" the frontier estimate) can be set using a simple rule-of-thumb, while the second bandwidth (controlling

the degree of smoothing of the initial frontier estimate) can be optimized using by iterating the bootstrap along the lines of Simar and Wilson (2001, 2004).

In Appendix A we extend the computationally efficient bootstrap method of Kneip et al. (2011) to estimate confidence intervals for directional distance functions. While the extension is straightforward, some details differ from the algorithm given by Kneip et al. due to the introduction of an arbitrary direction in (2.6). In addition, when the direction vectors d_x and d_y are strictly positive, the need for extrapolation as required by the Kneip et al. method can be avoided.

While the computationally efficient bootstrap described in Appendix A can be used to estimate confidence intervals for the directional efficiency of individual firms or points in Ψ, it is inappropriate for testing hypotheses about the structure of Ψ or its frontier. In such situations, the m out of n bootstrap analyzed by Simar and Wilson (2011) can be used for testing general hypotheses about Ψ and other features of the model.

The m out of n bootstrap is based on drawing, without replacement, $m < n$ times from the empirical distribution of the observed input-output pairs; Simar and Wilson (2011) show that ideas for choosing the subsample size m suggested by Politis et al. (2001) work well in finite samples, and prove that the method provides statistically consistent inference for testing hypotheses about the efficiency of individual points as well as for testing general hypotheses using statistics that are Borel functions of a set of distance functions. Results from Monte Carlo experiments reported by Simar and Wilson (2011) indicate that the method works well in finite samples, yielding tests with appropriate size and confidence interval estimates with coverage close to nominal levels.

The theoretical treatment in Simar and Wilson (2011) is in terms of input distance functions, but is easily extended to the directional distance functions considered here. Substituting the directional distance function defined in (2.6) and its VRS-DEA and FDH estimators for the input distance function and its corresponding estimators appearing in Section 4 of Simar and Wilson (2011), results similar to those obtained in Simar and Wilson are obtained since the asymptotic results obtained above in Section 3 are analogous to properties of the DEA estimator of the input-oriented distance function in (2.3).

Use of the m out of n bootstrap to estimate confidence intervals for $\delta(x, y \mid d_x, d_y, \Psi)$ corresponding to the point (x, y) is straightforward and analogous to the description in

Simar and Wilson (2011) regarding estimation of confidence intervals for the input-oriented efficiency measure. Consider an estimate $\widehat{\delta} = \widehat{\delta}_{\text{VRS}}(x, y \mid d_x, d_y, \mathcal{X}_n)$ of $\delta(x, y \mid d_x, d_y, \Psi)$, and corresponding bootstrap estimates $\widehat{\delta}_b^* = \widehat{\delta}(x, y \mid d_x, d_y, \mathcal{X}_m^*)$, $b = 1, \ldots, B$ based on bootstrap samples $\mathcal{X}_m^*$ of size $m < n$ obtained by drawing m times, independently, uniformly, and without replacement from $\mathcal{X}_n$. From Theorem 4.1, we have

$$\Pr\left(a_\alpha \leq n^{2/(p+q+1)}\left(\widehat{\delta} - \delta\right) \leq b_\alpha\right) = (1-\alpha) \tag{5.1}$$

for some a_α and b_α. Moreover, the results of Simar and Wilson (2011) adapted to the case of directional distance functions and Theorem 2.1 in Politis et al. (2001) ensure that

$$\Pr\left(a_\alpha \leq n^{2/(p+q+1)}\left(\widehat{\delta}^* - \widehat{\delta}\right) \leq b_\alpha\right) \approx (1-\alpha), \tag{5.2}$$

with the approximation improving as $B \to \infty$. Given $\widehat{\delta}$ and the B bootstrap values $\widehat{\delta}_b^*$, a_α and b_α can be estimated by the $\frac{\alpha}{2}$ and $(1 - \frac{\alpha}{2})$ percentiles of the set of values $\left\{m^{2/(p+q+1)}(\widehat{\delta}_b^* - \widehat{\delta})\right\}_{b=1}^{B}$; denote these estimates by $\widehat{a}_\alpha^*$ and $\widehat{b}_\alpha^*$, respectively. Substituting these estimates for a_α and b_α in (5.1) and then rearranging terms yields

$$\left(\widehat{\delta}_{\text{lo},\alpha},\ \widehat{\delta}_{\text{hi},\alpha}\right) = \left(\widehat{\delta} - n^{-2/(p+q+1)}\widehat{b}_\alpha^*,\ \widehat{\delta} - n^{-2/(p+q+1)}\widehat{a}_\alpha^*\right), \tag{5.3}$$

which provides an estimate of the $(1-\alpha) \times 100$-percent confidence interval for δ in (5.1).[7]

The only remaining issue is how to choose the sub-sample size m; this can be done using the data-driven method described in Simar and Wilson (2011). Consider a set $\mathbb{M} = \{m_j\}_{j=1}^{J}$ of sub-sample sizes where $m_1 < m_2 < \ldots < m_J$.

In the case of confidence intervals, estimated lower and upper bounds $(\widehat{c}_{\text{lo},j}(\alpha), \widehat{c}_{\text{hi},j}(\alpha))$ can be computed for each $m_j \in \mathbb{M}$. Then for each $j \in \{2,\ 3,\ (J-1)\}$, compute V_j as the sums of the standard deviations of $\{\widehat{c}_{\text{lo},j-1}(\alpha),\ \widehat{c}_{\text{lo},j}(\alpha),\ \widehat{c}_{\text{lo},j+1}(\alpha)\}$ and $\{\widehat{c}_{\text{hi},j-1}(\alpha),\ \widehat{c}_{\text{hi},j}(\alpha),\ \widehat{c}_{\text{hi},j+1}(\alpha)\}$; choose the subsample size and corresponding confidence interval estimate corresponding to $\min_j V_j$. In the case of hypothesis testing, we can proceed similarly by computing critical values $c_j(\alpha)$ corresponding to each $m_j \in \mathbb{M}$, and then for each $j \in \{2,\ 3,\ (J-1)\}$, compute V_j as the standard deviation of $\{c_{j-1}(\alpha),\ c_j(\alpha),\ c_{j+1}(\alpha)\}$.

[7] Simar and Wilson (2011) also discuss use of the m out of n bootstrap for testing hypotheses about model structure (e.g., returns to scale, convexity of Ψ) using input-oriented distance functions. It is straightforward to adapt the testing methods discussed by Simar and Wilson (2011) for use with directional distance functions.

Then choose the sub-sample size and corresponding critical value corresponding to $\min_j V_j$. In either case, V_j gives a measure of "volatility" of either the estimated confidence bounds or the estimated critical values. See Simar and Wilson (2011) for additional discussion.

6 Empirical Illustrations

Charnes et al. (1981, pp. 680–682) list 70 observations on 5 inputs and 3 outputs of schools examined in their study.[8] These data have been examined by Wilson (1993), Simar (2003), and others; although the number of dimensions (eight) is large for the number of observations (70) given the slow convergence rate ($n^{2/9}$) of the VRS-DEA estimator, the data serve as a useful, illustrative application that readers can replicate. We use the Charnes et al. (1981) data to illustrate the full-sample bootstrap method explained in Appendix A and adapted from Kneip et al. (2011).

Application of the full-sample bootstrap is straight-forward. For each observation, we take the direction vectors d_x, d_y to be the observed input-output vectors for the given observation; we then transform the data to the (Z, U)-space as described in Section 3, and employ the algorithm given in Appendix A within a golden section search to optimize the bandwidth h.[9] For each observation, we use $B = 2,000$ bootstrap replications.

Results are reported in Table 1, consisting of two panels with the one on the left giving results for observations 1–35 and the one on the right giving results for observations 36–70. In each panel, the column labeled "$\widehat{\delta}_i$" gives the estimate of the directional distance function defined in (2.6) for observations i obtained with the estimator in (3.5). The next two columns give the estimated 95-percent confidence interval, while the remaining columns give the optimized values of the bandwidths and finally the estimated size (i.e., one minus the estimated coverage) of the estimated confidence intervals. Size can be estimated because the full-sample bootstrap involves iteration to optimize the smoothing parameter h. Although we do not do so here, one could adjust the nominal size in order to optimize the achieved size of the estimated confidence intervals as described by Simar and Wilson (2001, 2004).

As discussed previously, the DEA estimator is biased downward. This is reflected in

[8] See Charnes et al. (1981) for a complete description of the data, including precise definitions of the inputs and outputs.

[9] See Press et al. (2007) for a description of the golden-section search method

the results given in Table 1 by the fact that the distance function estimates lie to the left of the confidence interval estimates. This is not surprising and is as it should be—the bootstrap that we have used incorporates an automatic correction for bias when estimating confidence intervals. The estimated sizes average 0.0681 over the 70 observations in the sample, ranging from 0.0235 to 0.1490. On average, the estimated sizes a are close to the nominal value of 0.05; some variation is to be expected due to the small sample size and the large dimensionality of the problem.

As a second example, we use sub-sampling methods to examine data on 129 aggressive-growth Mutual Funds; data were collected from Morningstar and updated May 2002. These same data have been examined by Simar and Vanhems (2010), and with additional mutual fund data by Daraio and Simar (2006, 2007).[10] In this setting, the traditional output Y is *total annual return*, expressed in percentage terms; we consider three inputs, namely *risk*, given by the standard deviation of returns (X_1), the *expense ratio* which measures of transaction costs (operating expenses and management fees, administrative fees, and all other asset-based costs) as a percentage of total assets (X_2), and X_3, the *turnover ratio* that measures the fund's trading activity (X_3; funds with higher turnover incur greater brokerage fees for affecting the trades). Annual returns can be either positive or negative; due to the nature of aggressive growth funds and the period covered by our data (January 1–December 31, 2001), most of the returns observed in our sample are negative.

We take the direction vectors d_x, d_y as the average values for the inputs (all are positive in our case) so that $d_x = (34.98, 155.19, 1.68)'$, and the average of absolute values of returns, yielding $d_y = 18.36$ Note that since the output is here univariate, the value of d_y does not matter so much (it is only a rescaling of the directional distance). Comparing the efficiency level along an average benchmark direction sounds meaningful, although from a theoretical view point, any other direction could be chosen.

Table 2 shows individual results for 20 funds randomly selected from the data.[11] For

[10] We are grateful to Cinzia Daraio for providing the data. A number of studies have applied efficiency and productivity measurement techniques to the problem of evaluating the performances of mutual funds. In these studies, risk (measured by volatility or variance) and average return on a fund or portfolio are analogous to inputs and outputs in models of production. The boundary of the attainable set of funds gives a benchmark relative to which the efficiency of a fund can be measured (see, for example, Briec et al., 2004 and Briec and Kerstens, 2009 for discussions of the relation between the hypothesis of the basic Markowitz model and production theory, and the usefulness of directional distances in this framework).

[11] To facilitate comparison, the selected units are the same as those selected in Simar and Vanhems (2010).

each selected observation, Table 2 shows FDH and VRS-DEA estimates of the directional distance functions in columns 2 and 3. In addition, in columns 4–6, results are shown for the bias-corrected version of the VRS-DEA estimator. The bias-corrected estimates shown in column 5 are computed using

$$\widehat{\delta}_{\mathrm{BC,DEA}} = \widehat{\delta}_{n,\mathrm{DEA}} - (m/n)^{2/(p+q+1)} \frac{1}{B} \sum_{b=1}^{B} (\widehat{\delta}^{*}_{b,\mathrm{DEA}} - \widehat{\delta}_{n,\mathrm{DEA}}) \tag{6.1}$$

where we adjust for the difference between the original sample size n and the sub-sample size $m < n$. For each observation, the sub-sample size m_{BC} appearing in column 4 is chosen to minimize the volatility of the bias-correction term, again along the lines of Simar and Wilson (2011). Column 6 in Table 2 gives the corresponding bootstrap estimates of the standard errors of the initial, uncorrected VRS-DEA estimates. Comparing the values in Columns 3 and 5 with the standard errors in column 6 reveals that the estimated bias is larger in every case than $0.25 \times \sigma(\widehat{\delta}_{\mathrm{DEA}})$. Employing the conservative rule given in Efron and Tibshirani (1993), the results suggest that the bias corrected estimate should be used in favor of the uncorrected estimate.

The last three columns of Table 2 show results for confidence intervals estimated by the sub-sampling bootstrap. The sub-sample sizes m_{CI} used for constructing the individual interval estimates are given in column 7, and the estimated lower and upper bounds for 95-percent confidence intervals are given in the last two columns. The sub-sample sizes for estimating confidence intervals are sometimes different from the sub-sample sizes for bias-correction. Since the choice of sub-sample size is data-driven, this should not be surprising; in the case of confidence interval estimation, the sub-sample size m_{CI} is chosen to minimize the volatility of the estimated confidence bounds (see Simar and Wilson, 2011 for details and discussion), while the sub-sample size m_{BC} was chosen by minimizing the volatility of the bias-corrected estimate as discussed above. The width of the estimated confidence intervals gives an idea of the precision of the VRS-DEA estimates; while there is some variation in widths, in most cases the estimated intervals are rather narrow.

Using robust versions of directional distance estimators, Simar and Vanhems (2010) found evidence that fund #56 is an outlier in the sense that it does not lie close to other observations in the data. This is reflected in the confidence interval estimate for this observation; the estimated interval is wide—0 to 0.1309—reflecting the fact that there is little information in

the sample to inform us about the directional distance from this observation to the boundary of the production set. We know that the precision of the distance function estimates depends on the density of data points in a neighborhood of the frontier along the directional path $(-d_x, d_y)$; where the density is large, estimated confidence intervals are narrow, but where the density is sparse, estimated confidence intervals are wider. In contrast to the previous example with the school data from Charnes et al. (1981), here the estimated lower bounds of the confidence intervals typically coincide with the initial VRS-DEA estimate. In the previous example, the lower bounds were to the right of the initial distance function estimate. This difference is due to the fact that here, the convergence rate for the VRS-DEA estimator is $n^{2/5}$, whereas in the previous example, the convergence rate was $n^{2/9}$. In addition, the sample size is larger in this example than in the previous example. Here, this bias is less than in the previous example, though perhaps still substantial.

7 Conclusions

This paper, examines the non-parametric DEA estimator of directional distances. Directional distances generalize the standard input or output oriented radial distances considered by Debreu (1951), Farrell (1957), Shephard (1953, 1970), and others, and consequently allow for more general analyses of production. In addition, directional distance functions and their estimators are able to accommodate negative quantities of inputs or outputs, whereas the traditional measures do not. We extend existing results on asymptotic convergence properties of radial DEA estimators, obtained by Kneip et al. (2008), to the case of directional distances. This allows adaptation of bootstrap procedures developed by Kneip et al. (2008, 2011) and Simar and Wilson (2011) for radial distance functions to the case of directional distances. We provide empirical examples to illustrate the use of both full-sample and sub-sampling bootstrap methods for making inference and testing hypotheses about model structure. Our illustration of the test of convexity in our second example is easily extended to tests of returns to scale as in Simar and Wilson (2011) and other model features.

A Appendix: Bootstrap For Confidence Intervals

Kneip et al. (2011) developed a consistent bootstrap for making inferences about input- or output-oriented efficiency measures that avoids much of the computational burden of the double-smooth bootstrap algorithm developed by Kneip et al. (2008). The main idea underlying the Kneip et al. (2011) bootstrap method is to use a naive bootstrap for observations that are "far" from the initial DEA estimate of the production frontier, while drawing from a smooth, uniform distribution for observations that lie "near" the DEA frontier. The method requires two bandwidth parameters, h_1 and h_2, which can be optimized using data-driven methods as in Kneip et al. (2011). The bandwidth h_2 controls the degree of smoothing of the initial frontier estimate, while the bandwidth h_1 determines the size of the neighborhood near the initial frontier estimate. With a few changes, the idea can be extended to the case of directional distances; given values for the bandwidth parameters, the following algorithm is analogous to Algorithm #2 appearing in Kneip et al. (2011):

Input: $\mathcal{X}_n,\ h_1,\ h_2,\ w_0,\ d_0,\ \alpha$

Output: $\widehat{\delta}_{\text{VRS}}(x, y \mid \text{d}_x, d_y, \mathcal{X}_n)$, $(\widehat{\delta}_{\text{lo},\alpha}, \widehat{\delta}_{\text{hi},\alpha})$, $\widehat{\alpha}(h_2)$

---

[1] Using the transformation defined in (4.1), transform each $(X_i, Y_i) \in \mathcal{X}_n$ to (Z_i, U_i) to form the set $\mathcal{Z}_n$.

[2] Compute $\widehat{\delta}_0 = \widehat{\delta}_{\text{VRS}}(x_0, y_0 \mid d_x, d_y, \mathcal{X}_n) = \frac{\widehat{\phi}_{\text{VRS}}(0|\mathcal{Z}_n, w_0)}{||d_0||}$, and $\widehat{\delta}_i = \frac{\left(\widehat{\phi}_{\text{VRS}}(Z_i|\mathcal{Z}_n, w_0) - U_i\right)}{||d_0||}$ $\forall\ i = 1,\ \ldots,\ n$.

[3] Set $h_1 = 4\widehat{\delta}_{\text{med}} n^{-2/(3(p+q+1))}$ where $\widehat{\delta}_{\text{med}}$ denotes the median of the values $\widehat{\delta}_1,\ \widehat{\delta}_2,\ \ldots, \widehat{\delta}_n$ computed in step [2].

[4] Compute the smoothed frontier points (Z_i, U_i^{∂}) for each $i = 1,\ \ldots,\ n$ where $U_i^{\partial} = \widehat{\phi}_{\text{VRS}}(0 \mid \mathcal{Z}_n, w_0) + h_2^2 \left[\widehat{\phi}_{\text{VRS}}(h_2^{-1} Z_i \mid \mathcal{Z}_n, w_0) - \widehat{\phi}_{\text{VRS}}(0 \mid \mathcal{Z}_n, w_0)\right]$.

[5] Set $k = 0$, $\mathcal{B} = \varnothing$.

[6] Loop over steps [5.1]–[5.9] B_1 times:

[6.1] Draw independently, uniformly, and with replacement from the set of integers $\{i\}_{i=1}^n$ n times to create a set of labels $\mathcal{J} = \{j_i\}_{i=1}^n$.

[6.2] For each $i = 1, \ldots, n$, set $Z_i^* = Z_{j_i}$ and

$$U_i^* = \begin{cases} U_{j_i}^{\partial} - \widehat{\delta}_{j_i} & \text{if } \widehat{\delta}_{j_i} > h_1; \\ U_{j_i}^{\partial} - \xi_{j_i}^* & \text{otherwise,} \end{cases}$$

where $\xi_{j_i}^*$ is a random, independent draw from a uniform distribution on the interval $[0, h_1]$, to construct a bootstrap sample $\mathcal{Z}_n^* = \{(Z_i^*, U_i^*)\}_{i=1}^n$.

[6.3] Compute $\widehat{\delta}_0^* = \frac{\widehat{\phi}_{\text{VRS}}(0|\mathcal{Z}_n^*, w_0)}{||d_0||}$ and add $\widehat{\delta}_0^*$ to the set $\mathcal{B}_{h_2}$.

[6.4] Analogous to step [2], compute $\widehat{\delta}_i^* = \frac{(\widehat{\phi}_{\text{VRS}}(Z_i^*|\mathcal{Z}_n^*, w_0) - U_i^*)}{||d_0||}$ $\forall\ i = 1, , \ldots, n$.

[6.5] Compute (smoothed) frontier points $(Z_i^*, U_i^{\partial *})$ where $U_i^{\partial *} = \widehat{\phi}_{\text{VRS}}(0 \mid \mathcal{Z}_n^*, w_0) + h_2^2 \left[\widehat{\phi}_{\text{VRS}}(h_2^{-1} Z_i^* \mid \mathcal{Z}_n^*, w_0) - \widehat{\phi}_{\text{VRS}}(0 \mid \mathcal{Z}_n^*, w_0)\right]$ $\forall\ i = 1, \ldots, n$.

[6.6] Set $\mathcal{B}_h^* = \varnothing$.

[6.7] Loop over steps [6.7.1]–[6.7.3] B_2 times:

[6.7.1] Draw independently, uniformly, and with replacement from the set of integers $\{i\}_{i=1}^n$ n times to create a set of labels $\mathcal{J}^* = \{j_i^*\}_{i=1}^n$.

[6.7.2] For each $i = 1, \ldots, n$, set $Z_i^{**} = Z_{j_i}^*$ and

$$U_i^{**} = \begin{cases} U_{j_i}^{\partial *} - \widehat{\delta}_{j_i^*}^* & \text{if} \widehat{\delta}_{j_i^*}^* > h_1; \\ U_{j_i}^{\partial *} - \xi_{j_i^*}^{**} & \text{otherwise,} \end{cases}$$

where $\xi_{j_i^*}^{**}$ is a random, independent draw from a uniform distribution on the interval $[0, h_1]$, to construct a bootstrap sample $\mathcal{Z}_n^{**} = \{(Z_i^{**}, U_i^{**})\}_{i=1}^n$.

6.7.3] Compute $\widehat{\delta}_0^{**} = \frac{\widehat{\phi}_{\text{VRS}}(0|\mathcal{Z}_n^{**}, w_0)}{||d_0||}$; add $\widehat{\delta}_0^{**}$ to the set $\mathcal{B}_{h_2}^*$.

[6.8] Use the estimate $\widehat{\delta}_0^*$ computed in step [6.3] and the set $\mathcal{B}_{h_2}^* = \{\widehat{\delta}_{0,\ell}^{**}\}_{\ell=1}^{B_2}$ of bootstrap values to estimate a $(1-\alpha) \times 100$-percent confidence interval $\left[\widehat{c}_{\text{lo},\alpha(h_2)}^*,\ \widehat{c}_{\text{hi},\alpha(h_2)}^*\right]$ for $\widehat{\delta}_0$.

[6.9] If $\widehat{\delta}_0 \in \left[\widehat{c}_{\text{lo},\alpha(h_2)}^*,\ \widehat{c}_{\text{hi},\alpha(h_2)}^*\right]$ then increment k by 1.

[7] Use the estimate $\widehat{\delta}_0$ computed in step [2] and the set $\mathcal{B}_{h_2} = \{\widehat{\delta}_{0,\ell}^*\}_{\ell=1}^{B_1}$ of bootstrap values to estimate a $(1-\alpha) \times 100$-percent confidence interval $\left[\widehat{c}_{\text{lo},\alpha(h_2)},\ \widehat{c}_{\text{hi},\alpha(h_2)}\right]$ for δ_0.

[8] Compute $\widehat{\alpha}(h_2) = 1 - kB_2^{-1}$, the estimated size of the interval computed in step [7].

Kneip et al. (2011) discuss the need for extrapolation in their algorithm; extrapolation is needed when smoothing of the initial frontier estimates causes some observations to be shifted above the convex hull of the sample observations when working in the input direction (or to be shifted to the left of the convex hull of the sample observations when working in the output direction). With directional distances, however, provided each element of the direction vectors d_x and D_y are strictly positive, no extrapolation is needed. The smoothing in step [3] is certain to maintain convexity of the frontier estimate; see Kneip et al. (2011) for discussion.

The computations in steps [1]–[5] of the above algorithm can be done prior to the bootstrap loop beginning in step [6]; the computations need only be done once. As noted above, the bandwidth h_1 that is computed in step [3] controls the size of the neighborhood near the frontier estimate. From the discussion in Section 4.5 of Kneip et al. (2011), this bandwidth must be of order $O\left(n^{-2/3(p+1+1)}\right)$. Reflecting the δ_i computed in step [2] around zero yields a set of $(2n)$ points whose density must be symmetric around zero. As in Kneip et al. (2011), the bandwidth h_1 can be optimized using the normal reference rule of Freedman and Diaconis (1981) for selecting bin-widths in histogram estimators of probability density functions. Their rule sets bin-widths for a histogram estimator of the density of the set of $2n$ values $\mathcal{D} = \{\widehat{\delta}_i, -\widehat{\delta}_i\}_{i=1}^{n}$ to $2(\text{IQ})(2n)^{-1/3}$, where IQ denotes the inter-quartile range of the values in $\mathcal{D}$. Noting that the interquartile range of the values in $\mathcal{D}$ is simply the median of the n values $\widehat{\delta}_1,\ \widehat{\delta}_2,\ \ldots,\ \widehat{\delta}_n$, denoted by $\widehat{\delta}_{\text{med}}$, and multiplying the Freedman and Diaconis by $n^{-2/3(p+q+1)}/n^{-1/3}$ to obtain the correct order, gives the value for h_1 appearing in step [4] above.

As discussed in Kneip et al. (2011), the above algorithm can be embedded in a golden-section search algorithm (Kiefer, 1953) in order to optimize the value of the bandwidth h_2 that controls the degree of smoothing of the frontier estimate.

References

Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen (1999), *Lapack User's Guide, Third Edition*, Philadelphia, PA: Society for Industrial and Applied Mathematics.

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

Briec, W. and K. Kerstens (2009), Multi-horizon markowitz portfolio performance appraisals: a general approach, *Omega, International Journal of Management Science* 37, 50–62.

Briec, W., K. Kerstens, and J. Lesourd (2004), Single-period markowitz portfolio selection, performance gauging and duality: a variation on the luenberger shortage function, *Journal of Optimization Theory and Applications* 120, 1–27.

Chambers, R. G., Y. Chung, and R. Färe (1998), Profit, directional distance functions, and nerlovian efficiency, *Journal of Optimization Theory and Applications* 98, 351–364.

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

— (1981), Evaluating program and managerial efficiency: An application of data envelopment analysis to program follow through, *Management Science* 27, 668–697.

Daraio, C. and L. Simar (2006), A robust nonparametric approach to evaluate and explain the performance of mutual funds, *European Journal of Operational Research* Forthcoming.

— (2007), *Advanced Robust and Nonparametric Methods in Efficiency Analysis*, New York: Springer Science+Business Media, LLC.

Daraio, C., L. Simar, and P. W. Wilson (2010), Testing whether two-stage estimation is meaningful in non-parametric models of production. Discussion paper #1031, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

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

Färe, R. and S. Grosskopf (2000), Theory and application of directional distance functions, *Journal of Productivity Analysis* 13, 93–103.

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.

Färe, R., S. Grosskopf, D. W. Noh, and W. Weber (2005), Characteristics of a polluting technology: Theory and practice, *Journal of Econometrics* 126, 469–492.

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

Freedman, D. and P. Diaconis (1981), On the histogram as a density estimator: L_2 theory, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gegiete* 57, 453–476.

Fukuyama, H. (2003), Scale characterizations in a DEA directional technology distance function framework, *Europea Journal of Operations Research* 144, 108–127.

Gattoufi, S., M. Oral, and A. Reisman (2004), Data envelopment analysis literature: A bibliography update (1951–2001), *Socio-Economic Planning Sciences* 38, 159–229.

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

Kiefer, J. (1953), Sequential minimax search for a maximum, *Proceedings of the American Mathematical Society* 4, 502–506.

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

— (2011), A computationally efficient, consistent bootstrap for inference with non-parametric DEA estimators, *Computational Economics* Forthcoming.

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

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

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

Park, K. H. and W. L. Weber (2006a), A note on efficiency and productivity growth in the Korean banking industry, 1992–2002, *Journal of Banking and Finance* 30, 2371–2386.

— (2006b), Profitability of Korean banks: Test of market structure versus efficient structure, *Journal of Economics and Business* 58, 222–239.

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

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (2007), *Numerical Recipes: The Art of Scientific Computing*, Cambridge: Cambridge University Press.

Ricazo-Tadeo, A. J., E. Reig-Martinez, and F. Hernández-Sancho (2005), Directional distance functions and environmental regulation, *Resource and Energy Economics* 27, 131–142.

Schubert, T. and L. Simar (2011), Innovation and export activities in the german mechanical engineering sector: An application of testing restrictions in production analysis, *Journal of Productivity Analysis* Forthcoming.

Shephard, R. W. (1953), *Cost and Production Functions*, Princeton: Princeton University Press.

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

Silva-Portela, M. C. A., E. Thanassoulis, and G. Simpson (2004), Negative data in DEA: A directional distance approach applied to bank branches, *Journal of the Opertations Research Society* 55, 1111–1121.

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

Simar, L. and A. Vanhems (2010), Probabilistic characterization of directional distances and their robust versions. Discussion paper #1040, Institut de Statistique, Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

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

— (2001), Aplicación del 'bootstrap' para estimadores d. e. a., in A. Álvarez, ed., *La Medición de la Eficiencia y la Productividad*, Madrid: Pirámide. Translation of Performance of the Bootstrap for DEA Estimators and Iterating the Principle, Discussion Paper No. 0002, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

— (2004), Performance of the bootstrap for DEA estimators and iterating the principle, in W. W. Cooper, L. M. Seiford, and J. Zhu, eds., *Handbook on Data Envelopment Analysis*, chapter 10, New York: Kluwer Academic Publishers, Inc., pp. 265–298.

— (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.

— (2011), Inference by the m out of n bootstrap in nonparametric frontier models, *Journal of Productivity Analysis* Forthcoming.

Wheelock, D. C. 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.

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

— (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. Forthcoming.

Table 1: Directional Distance Function Estimates for Charnes et al. (1981) Data with Estimated 95-percent Confidence Intervals

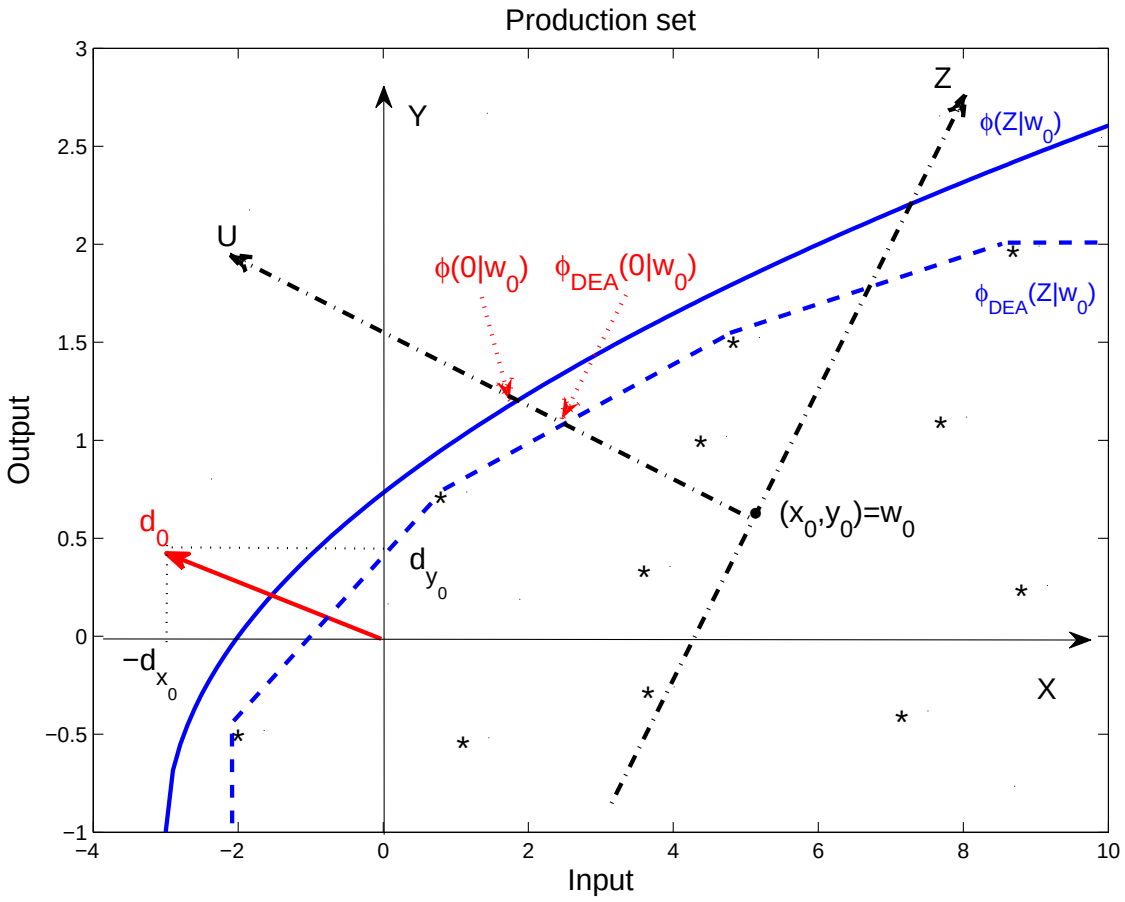
i	$\widehat{\delta}_i$	$\widehat{\delta}_{\mathrm{lo},\alpha}$	$\widehat{\delta}_{\mathrm{hi},\alpha}$	$\widehat{h}_1$	$\widehat{h}_2$	$1-\widehat{\alpha}(\widehat{h}_2)$	i	$\widehat{\delta}_i$	$\widehat{\delta}_{\mathrm{lo},\alpha}$	$\widehat{\delta}_{\mathrm{hi},\alpha}$	$\widehat{h}_1$	$\widehat{h}_2$	$1-\widehat{\alpha}(\widehat{h}_2)$
1	0.0174	0.0213	0.1686	0.5365	0.9895	0.0695	36	0.1172	0.1278	0.3677	0.4990	0.4204	0.0610
2	0.0519	0.0585	0.1950	0.4571	0.4715	0.0635	37	0.0881	0.0971	0.2498	0.4575	0.2681	0.0650
3	0.0334	0.0393	0.1223	0.4981	0.9895	0.1155	38	0.0000	0.0185	0.6667	0.5087	0.6709	0.0560
4	0.0514	0.0613	0.2136	0.5157	0.2090	0.0575	39	0.0316	0.0370	0.1460	0.4407	0.9554	0.0960
5	0.0000	0.0276	0.7532	0.6429	0.7560	0.0565	40	0.0258	0.0343	0.2333	0.5512	0.3929	0.1005
6	0.0485	0.0611	0.1983	0.4630	0.9895	0.0415	41	0.0241	0.0302	0.1279	0.4691	0.2316	0.0905
7	0.0568	0.0662	0.2377	0.5115	0.2591	0.0835	42	0.0267	0.0362	0.2505	0.4795	0.6563	0.0440
8	0.0497	0.0538	0.2143	0.4218	0.9895	0.0610	43	0.0728	0.0774	0.1760	0.4355	0.9895	0.0765
9	0.0752	0.0813	0.1337	0.4684	0.9895	0.0560	44	0.0000	0.0027	0.4341	0.4285	0.4950	0.0235
10	0.0289	0.0325	0.1352	0.4993	0.9554	0.0785	45	0.0000	0.0147	0.8095	0.4506	0.6563	0.0685
11	0.0000	0.0039	0.1748	0.4901	0.6879	0.0580	46	0.0451	0.0503	0.2500	0.5193	0.2681	0.0750
12	0.0000	0.0058	0.1659	0.5258	0.4099	0.0490	47	0.0000	0.0054	0.1526	0.4625	0.8282	0.1490
13	0.0731	0.0804	0.1690	0.4740	0.2316	0.0915	48	0.0000	0.0147	0.9011	0.5208	0.6328	0.0680
14	0.0062	0.0235	0.3744	0.5621	0.9895	0.0595	49	0.0000	0.0091	0.3394	0.4928	0.6118	0.0425
15	0.0000	0.0081	0.5088	0.4514	0.9108	0.0790	50	0.0212	0.0256	0.1244	0.5948	0.9895	0.0930
16	0.0255	0.0293	0.1964	0.4282	0.9895	0.1155	51	0.0417	0.0556	0.3882	0.6609	0.5882	0.0600
17	0.0000	0.0138	0.4946	0.5344	0.6879	0.0490	52	0.0000	0.0055	0.1828	0.5217	0.6879	0.0660
18	0.0000	0.0074	0.2082	0.5598	0.4545	0.0480	53	0.0690	0.0779	0.2453	0.5133	0.2762	0.0780
19	0.0241	0.0276	0.1117	0.4715	0.9895	0.0685	54	0.0000	0.0044	0.2666	0.5747	0.6879	0.0390
20	0.0000	0.0072	0.2010	0.4877	0.4124	0.0500	55	0.0003	0.0076	0.2049	0.4999	0.4294	0.0510
21	0.0000	0.0043	0.1889	0.5175	0.4991	0.0535	56	0.0000	0.0100	0.5297	0.4492	0.4991	0.0365
22	0.0000	0.0093	0.1239	0.4807	0.9895	0.0425	57	0.0376	0.0443	0.1452	0.5592	0.9895	0.0945
23	0.0126	0.0166	0.1003	0.4688	0.9895	0.0735	58	0.0000	0.0093	0.3188	0.5149	0.3127	0.0525
24	0.0000	0.0109	0.2278	0.5217	0.3548	0.0365	59	0.0000	0.0014	0.7677	0.5153	0.2681	0.0555
25	0.0107	0.0173	0.1408	0.4774	0.2381	0.0710	60	0.0098	0.0223	0.1673	0.5044	0.2381	0.0540
26	0.0294	0.0350	0.1323	0.4732	0.3403	0.1365	61	0.0627	0.0823	0.6050	0.6293	0.5096	0.0500
27	0.0000	0.0054	0.1671	0.4703	0.3548	0.0555	62	0.0000	0.0220	1.1583	0.4697	0.4950	0.0405
28	0.0054	0.0155	0.1524	0.4649	0.9895	0.0495	63	0.0187	0.0367	0.3020	0.5650	0.3022	0.0650
29	0.0784	0.1006	0.5573	0.5725	0.7771	0.0520	64	0.0358	0.0437	0.1634	0.4645	0.4399	0.1475
30	0.0560	0.0665	0.2521	0.4511	0.2721	0.0740	65	0.0129	0.0317	0.2734	0.5184	0.2251	0.0520
31	0.0884	0.0973	0.2491	0.4987	0.2551	0.0650	66	0.0330	0.0415	0.5308	0.7837	0.3022	0.0900
32	0.0000	0.0304	0.8673	0.6240	0.6433	0.0585	67	0.0273	0.0338	0.0926	0.6278	0.9895	0.0615
33	0.0243	0.0279	0.1220	0.4757	0.9895	0.0665	68	0.0000	0.0058	0.3256	0.4296	0.7220	0.0545
34	0.0751	0.0811	0.1674	0.4713	0.6879	0.1165	69	0.0000	0.0253	1.2876	0.7329	0.5266	0.0525
35	0.0000	0.0045	0.1583	0.4474	0.9173	0.0900	70	0.0181	0.0310	0.3116	0.5739	0.4820	0.0645

Table 2: FDH and DEA estimators of directional distance for 20 Mutual Funds

Units	$\widehat{\delta}_{\text{FDH}}$	$\widehat{\delta}_{\text{DEA}}$	m_{BC}	$\widehat{\delta}_{\text{BC,DEA}}$	$\sigma(\widehat{\delta}_{\text{DEA}})$	m_{CI}	$\widehat{\delta}_{\text{lo},\alpha}$	$\widehat{\delta}_{\text{hi},\alpha}$
3	0.0000	0.1097	62	0.1526	0.0363	122	0.1097	0.1472
99	0.5026	0.5290	62	0.5910	0.0594	76	0.5290	0.6968
107	0.1611	0.1711	34	0.2372	0.0478	122	0.1711	0.2007
39	0.4317	0.4627	80	0.5020	0.0675	122	0.4627	0.5059
51	0.1368	0.1867	80	0.2208	0.0532	40	0.1867	0.3537
121	0.2771	0.3205	76	0.3551	0.0441	122	0.3205	0.3534
122	0.0000	0.1020	62	0.1697	0.0566	100	0.1020	0.2018
15	0.1546	0.4047	46	0.4877	0.0476	84	0.4047	0.5564
123	0.0000	0.0745	62	0.1245	0.0297	122	0.0745	0.1389
28	0.2706	0.2946	106	0.3160	0.0410	112	0.2946	0.4085
65	0.3222	0.3372	90	0.3567	0.0322	74	0.3372	0.4818
56	0.0000	0.0000	80	0.0618	0.0866	122	0.0000	0.1309
115	0.3568	0.4898	62	0.5567	0.0479	120	0.4898	0.5645
27	0.4183	0.4224	106	0.4471	0.0531	94	0.4224	0.6067
6	0.3628	0.3706	106	0.4044	0.0714	88	0.3706	0.5861
31	0.0000	0.1182	64	0.1627	0.0393	100	0.1182	0.2215
61	0.4011	0.4081	106	0.4279	0.0410	116	0.4081	0.5130
45	0.4059	0.4744	120	0.4759	0.0053	122	0.4744	0.4827
91	0.7256	0.7283	90	0.7449	0.0455	122	0.7283	0.7337
129	0.2821	0.3571	106	0.3885	0.0627	96	0.3571	0.5299

NOTE: The value of m_{BC} gives the subsample size for bias correction and m_{CI} is for building the 95% confidence intervals.

Figure 1: Transformation from (x, y)-space to (z, u)-space



NOTE: The boundaries of the attainable set and its DEA estimate are indicated by the smooth, solid curve and the dashed, piece-wise linear curve, respectively. Data points are shown by asterisks (*).